Are Random Drift and Natural Selection Conceptually Distinct?

Roberta L. Millstein
Department of Philosophy
California State University, Hayward
rmillstein@csuhayward.edu

Abstract: The latter half of the twentieth century has been marked by debates in

evolutionary biology over the relative significance of natural selection and random drift:

the so-called "neutralist/selectionist" debates. Yet John Beatty has argued that it is

difficult, if not impossible, to distinguish the concept of random drift from the concept of

natural selection, a claim that has been accepted by many philosophers of biology. If this

claim is correct, then the neutralist/selectionist debates seem at best futile, and at worst,

meaningless. I reexamine the issues that Beatty raises, and argue that random drift and

natural selection, conceived as processes, can be distinguished from one another.

1. Introduction.

The role of random drift in twentieth century evolutionary biology has been a turbulent

one. Biologists such as Wright and Fisher disagreed over the relative importance of

random drift in the evolutionary process, with Fisher arguing that Wright's theory

assigned too great a role to random drift. The debate continues today, with "neutralists"

such as Kimura and Crow claiming an even larger role for random drift than Wright did,

while "selectionists" such as Ernst Mayr and Douglas Futuyma (see especially Mayr

1983 and Futuyma 1988) remain steadfast to the idea that it is natural selection that plays

the preeminent role in phenotypic evolution. According to the selectionist camp, the role

of random drift is at best a minor one. As Beatty (1984) argues, this debate has taken

place in specific contexts (over whether, for example, natural selection or random drift is

more prevalent in a particular population) as well as more general contexts (over whether

natural selection or random drift is more prevalent as a whole).

   With all this contention concerning the relative importance of random drift and

natural selection within evolution, one would hope that the concepts of random drift and

natural selection could be defined and distinguished clearly. After all, if one is to argue

that random drift is very prevalent in evolution, and natural selection less so (or vice

versa), it had better be the case that instances of natural selection can be distinguished

from instances of random drift – and in order to do *that*, the concepts must be distinct.

Even (or perhaps, especially) if one takes a non-empirical, theoretical approach to settling

this debate, it is still necessary that one be able to distinguish the concept of random drift

from the concept of natural selection. However, Beatty (1984) argues that it is not easy –

perhaps even impossible – to distinguish the two concepts. But if that is true, then it

seems as though the neutralist/selectionist debate has just been a tempest in a teapot, and

that participants in the debate have been arguing about nothing.

   Perhaps they have been. However, it is incumbent upon philosophers of biology to

try to clarify debates within biology such as these, if possible. It is with this motivation

that I seek to reexamine the questions that the debate presupposes. What is random drift?

Can it be distinguished from natural selection? In addressing these questions I will not,

however, take a stance on the neutralist/selectionist debate itself, an issue best left to

biologists.

   First, I characterize random drift and natural selection, both as processes and as

outcomes. Then I turn to the question of whether random drift can be distinguished from

natural selection. Using the characterization of random drift and natural selection that I

developed earlier, I argue that the concepts of natural selection and random drift are

distinct as processes, if not as outcomes. Lastly, I consider an alternative characterization

of natural selection and random drift, due to Brandon and Carson (1996), that implies the

same conclusion I reach but for different reasons. I argue that the account I have

presented is the preferable one.

2. Random Drift and Natural Selection: Processes and Outcomes.

The term 'random drift' actually refers to a number of distinct processes, not just one

kind of process, although the different processes have characteristics in common. Beatty

(1992) isolates four kinds of random drift: indiscriminate parent sampling, indiscriminate

gamete sampling,[1] the founder effect, and fluctuations in the rates of evolutionary

processes (selection, migration, and mutation). The founder effect is similar to another

phenomenon, known as the bottleneck effect. Another kind of random drift is the random

assortment of genes into gametes (Dodson and Dodson 1985, 229). Brandon and Carson

(1996) claim that Brandon (1990) identifies yet another "drift-like phenomenon." This

leaves potentially at least seven different kinds of random drift. I say "potentially"

because there has been some controversy as to which processes should be considered random drift (Beatty 1992). Indeed, different authors will describe different subsets of this list as constituting the different kinds of random drift processes. For reasons of space, I will discuss only two of these processes: indiscriminate parent sampling and the bottleneck effect. The discussion of different kinds of random drift illustrates the general characteristics of random drift and focuses attention on the biological basis for random drift. I will also contrast random drift with natural selection in relatively straightforward situations, saving the more complex cases for later discussion.

Parent sampling, according to Beatty, is "*the process of determining which organisms of one generation will be parents of the next, and how many offspring each parent will have*" (1984, 188; italics in original). According to Beatty, this parent sampling can be discriminate, that is, with regard to fitness differences, or indiscriminate, that is, without regard to fitness differences (1984, 189). Discriminate parent sampling is generally considered natural selection; indiscriminate parent sampling is random drift.

But what is sampling "with regard to fitness differences" and sampling "without regard to fitness differences"? A hypothetical example will help to clarify. Suppose a population of gray and brown squirrels shares its habitat with a colorblind predator. The gray squirrels die in greater numbers than the brown squirrels – it just so happens, say, that more of the gray squirrels are unlucky enough to come into contact with the predator. Consequently, the frequency of gray squirrels in the next generation is reduced. More brown squirrels are "sampled" (i.e., survive to reproduce) than gray squirrels, but it is without regard to any differences in fitness between the gray squirrels and the brown

squirrels. In other words, the physical differences between the two types of squirrels do

not play a *causal* role in the differences in reproductive success.[2] The change in

frequency in this squirrel population is due to the kind of random drift known as

indiscriminate parent sampling. This contrasts with alternative scenarios in which

physical differences *do* play a causal role in the differences in reproductive success.

Suppose, for example, that the predator was not colorblind, and that in this particular

environment the predator was able to see the gray squirrels more clearly than the brown

squirrels, causing greater survival of brown squirrels over gray. This would be a case of

discriminate parent sampling, or natural selection.

A related phenomenon is what is sometimes called the *bottleneck effect*. A bottleneck

occurs when there is a large reduction in the number of organisms in a population (as

occurs frequently in some populations that tend to fluctuate in size, perhaps in response

to fluctuating environment). When biologists speak of a bottleneck, they are generally

referring to an indiscriminate sampling process; physical differences between organisms

in the original population are causally irrelevant to differences in reproductive success

that produce the composition of the smaller population. For example, a severe drought

might devastate a population of artichokes, killing artichokes without regard to

differences in color, size, or shape. Logically, one could speak of a *discriminate*

bottleneck where there was a drastic reduction in population size *and* where the physical

differences of some organisms gave them a greater ability to survive and reproduce, but

such a phrase would be misleading. For example, it is hypothesized that cheetahs

underwent a severe bottleneck at some point during their evolutionary history. Suppose

that during this bottleneck, the cheetahs with longer legs had a greater ability to survive than the cheetahs with shorter legs. However, while such a discriminate sampling process (i.e., selection) might explain why longer-legged cheetahs were in the end more prevalent, it would not explain why there was an *overall* reduction in the number of cheetahs, both long-legged and short-legged. In the overall reduction of the population, organisms are killed indiscriminately. Thus, biologists seem to prefer to speak of such cases as natural selection *together with* random drift (as a bottleneck), rather than as a discriminate bottleneck. (We will discuss such an example below.) The bottleneck effect is quite similar to indiscriminate parent sampling, since both processes involve sampling of parents which is indiscriminate with respect to physical differences. The primary difference is that with the bottleneck effect, there is a drastic reduction in the size of the population, whereas with indiscriminate parent sampling, the size of the population may remain relatively constant from one generation to the next.

The two kinds of random drift that I have discussed here, indiscriminate parent sampling and indiscriminate bottlenecks, are processes that sample indiscriminately with respect to physical differences. That is, they are processes in which any physical differences are causally irrelevant in the sampling process. Analogously, the two kinds of natural selection I discussed, discriminate parent sampling and discriminate sampling during a bottleneck, are processes that sample discriminately with respect to physical characteristics, meaning that physical differences are causally relevant in the sampling process. Again, the difference has to do with whether population size remains relatively

constant from one generation to the next (parent sampling), or whether the population is reduced in size (bottleneck).

My suggestion that Beatty's conceptions of "discriminate sampling" and "indiscriminate sampling" should be spelled out in terms of causal relevance is similar to Hodge's (1987) position. Hodge asserts that a proper distinction between natural selection and random drift relies on *causation*: "… differential reproduction in selection is distinguished from any in drift by its causation; by contrast with drift, it is occurring because the physical property differences constituting the hereditary variation that is being differentially reproduced are not merely correlated with differences in reproduction – they are causally relevant to them" (1987, 251). Note that Hodge, as is commonly done, includes heredity as a necessary condition for both natural selection and random drift. I am in agreement with Hodge on this point as well, and all physical differences mentioned in subsequent discussion should be assumed to be heritable.

Beatty's discussion of random drift as "indiscriminate sampling" and natural selection as "discriminate sampling" focuses on random drift and natural selection as certain kinds of *processes*. However, it is important to note that while random drift is sometimes spoken of as a *process*, at other times it is spoken of as an *outcome*. Suppose there were a population of giraffes not undergoing mutation, migration, or selection – only random drift. Again, random drift is an indiscriminate sampling *process*, in which physical differences between organisms are causally irrelevant to their reproductive success. Thus, if the long-necked giraffes in the population become more prevalent than short-necked giraffes, it is by chance (and not because of their long necks). In an infinite (or

very large) population, we would expect the proportion of long-necked to short-necked

giraffes to remain relatively constant from generation to generation. However, we would

not expect the proportions to remain constant in a small population. To understand this,

imagine an urn filled with colored balls where balls are sampled without respect to color.

If a large sample of balls were taken, we would expect the frequencies of colored balls in

the sample to be very close to the frequencies in the urn. On the other hand, if we only

take a small sample of colored balls, our sample may very well have different proportions

of colored balls than the urn does. In the same way that the color difference between the

balls is causally irrelevant to which ball gets picked, in a population undergoing the

process of random drift, the physical differences between organisms are causally

irrelevant to differences in their reproductive success. So in large populations, as with

large urn samples, we would expect gene frequencies to be representative of the parent

generation, but in small populations, as with small urn samples, gene frequencies may or

may not be representative. Thus, when random drift occurs over a number of generations

in a small population, gene frequencies may fluctuate, or *drift*, randomly from generation

to generation.

However, when we speak of the drifting of gene frequencies in this way we are no

longer focusing on random drift as a certain kind of *process* (an indiscriminate sampling

process, analogous to sampling balls from an urn). We are considering random drift as a

certain kind of *outcome* (an outcome where gene frequencies fluctuate from generation to

generation, in the same way that repeated sampling of balls from an urn would fluctuate

from sample to sample). Similarly, we might speak of natural selection as a certain kind

of process (a discriminate sampling process), as well as a certain kind of outcome (an outcome where the fittest organisms have the greatest reproductive success).[3]

In distinguishing a process from its outcome, I mean to distinguish the kinds of changes that occur over time (the process) from the "ending" state that occurs at one point in time (the outcome). Of course, in a population undergoing evolution, designation of an ending state (or a beginning state, for that matter) at any particular point in time is arbitrary. To return to the indiscriminate parent sampling example discussed above, suppose that we (arbitrarily) choose the initial state of the population at a point in time in which the population is composed of equal numbers of gray and brown squirrels. The squirrels are exposed to the colorblind predator, causing the population (as described by the relative proportions of each color of squirrel) to change in a certain kind of way: the squirrels are sampled indiscriminately with respect to their color. In other words, the color difference between the squirrels is causally irrelevant to their survival. This is an instance of the process of random drift (more specifically, an instance of indiscriminate parent sampling). The end state, or outcome of this process (again, chosen at an arbitrary point in time), might be that population is composed of 51% brown squirrels and 49% gray squirrels. Alternatively, the outcome might be more dramatic – say, 60% brown squirrels and 40% gray squirrels. Thus, we may have different *outcomes* from the same kind of *process*; outcome is distinct from process. The case of the indiscriminate bottleneck is much the same, except that if there is a severe bottleneck (where the population is reduced to a very small size), the new population is that much more likely to experience an outcome which diverges widely from the original population. The

reduced population may carry only a small fraction of the variation of the original population.

In the discussion that follows, I will argue that this distinction between process and outcome is the key to solving the problem of distinguishing between random drift and natural selection.

3. Natural Selection and Random Drift: Conceptually Distinct or Not?

We have been discussing random drift and natural selection as processes that occur in isolation from one another. However, as Beatty (1984) notes, in any real, finite population where natural selection is occurring, random drift will be occurring as well. Thus, we need to be able to identify the different processes and outcomes as they occur in a population. Furthermore, as we discussed in section 1, there has been (and continues to be) a great deal of debate over the question of whether, in particular instances and in general, natural selection is more prevalent than random drift. For these reasons, it is important that the concept of random drift be distinguishable from the concept of natural selection. (How one distinguishes natural selection and random drift *empirically*, as opposed to *conceptually*, is a separate although not unrelated issue. This will be discussed further below.) But can they be so distinguished?

3.1. Beatty's "Chance and Natural Selection" Revisited.

According to Beatty, it is "conceptually difficult to distinguish natural selection from random drift" (1984, 196). Again, on Beatty's account, natural selection is *discriminate* sampling, or sampling with regard to fitness differences, whereas random drift is *indiscriminate* sampling, or sampling *without* regard to fitness differences (1984, 190-1). Yet Beatty maintains that these concepts are not conceptually distinct.

The potential inability to distinguish the concepts of natural selection and random drift is a problem for evolutionary biology given arguments over the relative importance of natural selection and random drift. How can natural selection and random drift be used as alternative explanations if the concepts of natural selection and random drift cannot be distinguished? Beatty suggests that some evolutionary changes are "to some extent, or in some sense, a matter of natural selection *and* to some extent, or in some sense, a matter of random drift" (1984, 196), acknowledging that, at least in some cases there *is* no hard and fast distinction between random drift and natural selection. This is not a solution to the problem he has raised, but rather a concession to the problem. Perhaps this is the best that we can do. However, I think it is important to try to make sense of biologists' empirical claims about whether random drift or natural selection is more prevalent in given population. In general, these claims are difficult to substantiate, but with ambiguous conceptual tools, they are impossible to substantiate. If Beatty is right, there will be some cases in which we cannot answer the empirical question because our concepts are not sufficiently distinct (this is different from, say, not being able to

answer the question because we can't obtain the relevant information). However, if the concepts are not distinct, then it seems as though *all* evolutionary explanations that invoke natural selection and/or random drift will suffer from vagueness, not just certain fuzzy cases. My motivation for the following discussion stems from a desire to avoid this unattractive conclusion.

As noted above, my definitions of natural selection and random drift are consistent with Beatty's and are essentially the same as those of Hodge (1987). Neither Hodge nor Beatty makes the explicit distinction between process and outcome that I made above, although both use the distinction implicitly. Citing Beatty (1984), Hodge states:

> … explicating selection by contrast with drift allows for – even encourages – the admission that in real life the ecology and genetics of butterflies and their predators may include some changes that are equally plausibly categorized as selection or drift. Once one considers such familiar complications as linked genes, correlated responses to selection, patchy environments, frequency dependent effects, habitat preferences, and so on, it is possible to think of scenarios for which it is impossible to draw a sharp line between [differential reproduction where physical property differences are causally irrelevant and differential reproduction where physical differences are causally relevant]. But here, as always, it is to be emphasized that such impossibilities do not nullify the rationales for making the conceptual distinction (1987, 253).

This passage may be interpreted as claiming that whereas, *empirically*, it may be difficult to classify some outcomes as random drift or natural selection, *conceptually*, the two are distinct. In what follows, I will revisit Beatty's arguments and use the process/outcome distinction to argue in support of Hodge's claim.

Beatty's conclusions are based on the following example of parent sampling. Consider a population of dark moths and light moths, both having the same color-sensitive predator. The moths inhabit a forest of light colored trees (40%) and dark colored trees (60%). The light colored moths are camouflaged against the light colored trees; the dark moths are camouflaged against the dark colored trees. Since there are more dark colored trees than light colored ones, the dark moths have a greater likelihood of landing on a tree that will prevent their detection by a predator. Thus, we would expect the dark colored moths to have greater reproductive success than the light colored moths.

However, as Beatty emphasizes, the expected reproductive success of each type of moth (dark or light) can be represented by a probability distribution where the probability of the given type for leaving each of the possible quantities of offspring is displayed. Now suppose that: 1) there is an overlap in the probability distributions of the dark and light moths, such that it is possible for the light moths to have greater *actual* reproductive success than the dark moths, even though the dark moths overall have a greater *expected* reproductive success; 2) the possible becomes actual, and the lighter moths *do* have greater reproductive success; and 3) the greater actual reproductive success of the lighter moths is due to the fact that, by chance, the dark moths more frequently landed on light

colored trees than dark colored trees, despite the fact that there was a greater proportion

of dark colored trees.

Given these three suppositions, Beatty asks: "Is the change in frequency of genes and

genotypes in question a matter of natural selection, or a matter of random drift? That is,

is the change in question the result of sampling discriminately or indiscriminately with

regard to fitness differences?" (1984, 195). Notice that in describing natural selection

and random drift in terms of *sampling*, the question that Beatty asks concerns natural

selection and random drift as *processes*. Let us now examine Beatty's exploration of

possible answers to this question, keeping in mind the distinction between process and

outcome.

First, Beatty states: "It is not easy to maintain that the sampling was entirely

indiscriminate with regard to differences in survival and reproductive ability" (1984,

195). Here Beatty is suggesting that we do not want to attribute the changes in the

population entirely to the random drift process. Beatty does not say why this is, but

presumably, it is because the light moths who perished did so on the basis of their

lightness; the lightness of these moths made them less fit than the darker moths in an

environment where the majority of trees were dark. In other words, the moths were

sampled discriminately, making the process of random drift implausible as a complete

explanation of the change. So far, so good.

Beatty then claims: "At least it is difficult to maintain that the death by predation of

conspicuously dark moths in this environment is indiscriminate sampling, whereas the

death of *conspicuously* light moths in the same environment is selection" (1984, 195). In

other words, it seems as though we would want to say that the light moths were sampled

discriminately with regard to their light color and thus, we would say that they were

subject to the process of natural selection. However, when we consider the dark moths

who died by predation, we might want to say that their deaths occurred indiscriminately

with regard to physical differences[4], since it is just by chance that so many of the dark

moths landed on light colored trees. If the dark moths were "sampled" without regard to

their darkness, we would want to say that they were undergoing the process of random

drift. But how could it be that the dark moths are undergoing the process of random drift,

while the light moths are undergoing the process of natural selection? After all, is there

anything that is different about the situations of the two kinds of moth? In both cases, the

colors of moths that died made the moths distinct against their backgrounds, leaving the

moths vulnerable to the predator. Why should one case be considered natural selection

and the other random drift? This answer seems implausible, and Beatty rightly rejects it.

He then suggests that "it is also difficult to maintain that selection alone is the basis of

the change. At least, it is difficult to maintain that the fittest were selected" (Beatty 1984,

195). Here lies the crux of the problem. Note that the former statement refers to the

process of natural selection: the *basis* of the change. However, the latter statement refers

to one possible *outcome* of natural selection: that the fittest were selected. Yet the

question that Beatty asked, "discriminate sampling or indiscriminate sampling?" was a

question about process, not outcome.

This conflation of process and outcome continues throughout the section. Beatty

thinks the problem in determining whether the moths exhibit random drift or natural

selection "is that it is difficult to distinguish between random drift on the one hand, and

the *improbable results of natural selection* on the other hand (1984, 195-6; italics in

original). Read "results" as "outcomes." Moreover, the conflation persists to Beatty's

conclusion that it is difficult to distinguish the concept of random drift from the concept

of natural selection:

> …it seems that we must say of some evolutionary changes that they are to some
>
> extent, or in some sense, a matter of natural selection *and* to some extent, or in
>
> some sense, a matter of random drift. And the reasons (one of the reasons) we must
>
> say this is that it is conceptually difficult to distinguish natural selection from
>
> random drift, especially where *improbable results of natural selection* are
>
> concerned (1984, 196; italics in original).

However, if we consider natural selection and random drift as processes only, the two

concepts *can* be distinguished in Beatty's example. On the one hand, we can identify the

occurrence of the process of natural selection. The color difference between the light and

the dark moths is causally relevant to the death of the light moths; the light moths died

because the predator could distinguish them more easily than the dark moths in an

environment where there are greater numbers of dark trees (similarly, the dark moths died

in an environment where the predator could distinguish them less easily). On the other

hand, there is no identifiable process where the physical differences in the population are

causally irrelevant to the differences in reproductive success. Thus, Beatty's example is

not a case of the processes of natural selection and random drift together; it is a case of the process of natural selection alone (with the qualification that since all real populations are finite, we cannot rule out the possibility of random drift entirely).

Now this might seem like an odd analysis for a population where dark moths are fitter than light moths, yet light moths are more reproductively successful. However, once we consider the reproductive success of the light moths, we are no longer analyzing processes; we are analyzing the outcome. So, some might claim that it is odd to say that an outcome where the fittest do not have the greatest reproductive success is an outcome of natural selection. (This seems to be Beatty's position.) This is a distinct claim from what I argued for above, an argument concerning processes.

Nonetheless, I would claim that Beatty's scenario *does* describe a natural selection outcome. This claim is plausible if we *take seriously the proposition that the theory of natural selection is probabilistic.*[5] If, as Beatty suggests, the fitness of an organism is represented by a probability distribution of possible outcomes, then certain outcomes are more likely than other outcomes, but why should the *less likely* outcomes be any less "representative" (to use Beatty's term) of natural selection than the *more likely* outcomes? All are predicted by natural selection; it is just that the predictions are of differing strengths.

More tellingly, consider the following: let's suppose that the population initially consisted of 50% light moths and 50% dark moths, but in the next generation (after dark moths unluckily land on too many light trees, and the surviving moths reproduce in equal numbers) the dark moths are 45% of the population and light moths are 55%. However,

since the most likely result (the mean or expected value) is 60% dark moths and 40% light moths, and since the probability distribution is more or less bell-shaped, the outcome of 75% dark moths and 25% light moths *has nearly the same probability* as the outcome of 45% dark moths and 55% light moths. Yet, with the latter possibility (75% dark moths and 25% light moths), where the fitter organisms *do* have the greater reproductive success, there is less difficulty in seeing that natural selection is the operative process. However, since both results are essentially equally likely, there is no reason to consider the one outcome (75% dark moths and 25% light moths) to be natural selection, and the other outcome (45% dark moths and 55% light moths) not to be.

There is a further aspect to consider in evaluating the claim that this is a population undergoing the process of natural selection. We are only looking at a change in the population from one generation of moths to another. Natural selection, as Darwin conceived it, is a process which occurs "in the course of many generations" ([1859] 1964, 114). In the short run, the unlikely can occur (the fittest organisms may be exposed to a disease, or encounter more predators), so that in the occasional generation, the fittest may not prevail. However, in the long run, we *would* expect the fitter organisms to be more reproductively successful. When we consider what would most likely happen to Beatty's moth population in succeeding generations, the claim that natural selection has occurred in the generation in question is more intuitive.

Although I have argued that Beatty's scenario is largely a case of natural selection, other scenarios may be analyzed as being cases of natural selection together with random drift.[6] Consider a population of pink and brown snails living in a region where most of

the surroundings are brown. In this environment, the brown snails are fitter (and better

adapted) than the pink snails, because the predator which feeds on the snails has a greater

ability to discern the pink snails, giving the pink snails a lower viability. Now suppose

that one summer there is a drought that causes both brown and pink snails alike to die in

great numbers. Let us further suppose that more brown snails are able to find moist

shelter and thus survive the drought than pink snails (and that the brownness of the brown

snails is causally unrelated to their greater ability to find shelter). Consequently, the next

generation of snails includes more brown snails than pink snails. Again, we can ask

whether this is a case of natural selection or random drift in order to determine whether

the concepts are distinct. I would argue that with respect to process, the concepts are

distinct.

In the snail population, there is a process in which the physical difference between the

snails is causally irrelevant to the difference in their reproductive success: this is the

(indiscriminate parent sampling) process whereby drought has produced an unusual

shortage of shelter, causing the death of pink and brown snails alike. Thus, the process of

random drift is occurring in the population. However, there is also a process in which the

heritable physical difference between the snails *is* causally relevant to the difference in

their reproductive success: some of the pink snails die because they are more easily

picked out by the predator in a largely brown environment. So, a discriminate parent

sampling process, natural selection, is occurring in the population as well. Thus, we can

distinguish the concepts of natural selection and random drift, at least with regard to

process, even when applied to a population in which both processes are occurring.

Note that my hypothetical snail example and the analysis I have given it have real-life counterparts. For example, in Rosemary and Peter Grant's study of *Geospiza conirostris* (large cactus finch), it was hypothesized that under dry conditions, "relatively long-billed birds should have been at a selective disadvantage because the foods they are best able to exploit, the fruits and flowers of *Opuntia*, become scarce. Relatively deep-beaked birds should have been at a selective advantage because the resources best exploited with a deep-beak, arthropods beneath bark and in the tissues of hardened, dead *Opuntia* pads, become relatively common" (Grant and Grant 1989, 220). Indeed, it was found that, "birds of both sexes with short bills survived best in the early phase of a long dry period when *Opuntia* flora resources declined. Birds with deep beaks were favored at a later stage when there was little food other than arthropods beneath bark and in *Opuntia* pads" (Grant and Grant 1989, 227).

To put these findings in the terms used above, at the beginning of the long dry period, physical differences (shorter bills) were causally relevant to differences in reproductive success. Towards the end of the dry period, the process changed: physical differences (deeper beaks) were causally relevant to differences in reproductive success. In other words, during a dry spell, there were two consecutive selective processes occurring, where there was first discriminate parent sampling for finches with shorter bills, and then for finches with longer beaks.

In addition to these selective processes, Grant and Grant speculate that there was probably random drift, in the form of a bottleneck, occurring in the population: "Conditions are right for occasional, extreme reductions in local neighborhood size.

Under these conditions random effects on allele frequencies may be very important. A

population or subpopulation passes through a bottleneck in which alleles are lost through

drift" (Grant and Grant 1989, 232). This quote implies both random drift processes and

outcomes, but for the moment, we will focus on the process alone. Fluctuations in the

environment caused periodic, drastic reductions in the size of the population – reductions

in which physical differences between finches were causally irrelevant to any differences

in reproductive success. Thus, natural selection and random drift, conceived as

processes, are distinguishable within the finch population.

What about random drift and natural selection conceived as outcomes – can they be

distinguished? If my snail example involves both natural selection and random drift

processes, then the question remains as to which outcome has prevailed. Here is where I

think we run into difficulty. One way that biologists commonly answer the question is to

say that natural selection has prevailed if $4N_e s \gg 1$, whereas random drift has prevailed if

$4Nes \ll 1$, where N_e is the effective population size and s is the selection coefficient

(Futuyma 1986, 173). In other words, when the effective population is large and/or the

selection coefficient is high, selection tends to prevail. When the effective population is

small and/or the selection coefficient is low, random drift tends to prevail. This rule of

thumb implies that there is a continuum of outcomes between random drift and natural

selection, suggesting that with regard to outcomes Beatty is correct: random drift and

natural selection are not distinct.[7] Furthermore, the rule of thumb leaves us with at least

two lingering concerns. First, since it is purely theoretical, it does not address the

question of which process *actually* prevailed to produce the outcome (if indeed such a

question can be answered). Second, it does not address the question as to what *extent*

each process has contributed to the overall outcome (and again, it is not clear that this

question can be answered).

Hodge (personal communication) has suggested that although it may be difficult to

determine whether one *particular* outcome – say, a 10% increase in type *A* over type *a* –

is due to random drift process or a natural selection process (since either random drift or

natural selection can produce such an outcome), it may be possible to determine whether

a *series* of outcomes is due to random drift or natural selection. More specifically, Hodge

suggests that a *consistent* and *repeatable* increase of *A* over *a* is evidence for selection, so

that the series of outcomes serves as evidence of the process.

Surely, this is correct, but it raises further questions. How consistent must the

increases be? How repeatable must they be? With regard to consistency, as we saw

above, sometimes natural selection can lead to an unexpected outcome, which in this case

would be a decrease in *A* over *a*. Clearly, we would want not one decrease to rule out

natural selection; so, for a series of outcomes, what percentage of outcomes may be

decreases for it still to be considered natural selection? With regard to repeatability, it is

not unlikely for random drift to produce a "streak of good luck" – an increase of one type

over another for a number of generations (Millstein 2000a). Apparent directionality does

not rule out random drift. So, while it is true that the *more* consistent and the *more*

repeatable that a series of outcomes is, the more likely it is to be natural selection, there

may be series of outcomes which cannot clearly be described as natural selection and

random drift.[8] All of which is to say that as outcomes, random drift and natural selection cannot always be distinguished (a point on which Hodge and I agree).

If the points I have argued for above can be granted, there are two possible conclusions that can be drawn. The obvious, straightforward conclusion is to accept that although random drift and natural selection can be distinguished when they are conceived as processes, they cannot be distinguished fully when they are conceived as outcomes. This conclusion accepts most of the current usages of the terms "natural selection" and "random drift" – that sometimes, when one says "natural selection" or "random drift", one refers to the processes, whereas other times, one refers to the outcomes. Clearly, however, the status quo paves the way for the kinds of problems discussed above, where process is confused with outcome.

An alternative, somewhat more normative conclusion proposes that this entire discussion of outcomes really should not be part of a discussion of concepts. Rather, outcomes – the *results* for particular populations undergoing the processes of natural selection and/or random drift – should more properly considered to be part of an *empirical* discussion, not a conceptual discussion. In other words, when we look at a particular population and ask whether a particular change was due to random drift or natural selection (or to what extent the change was due to random drift and to what extent the change was due to natural selection), we are really asking an empirical, not a conceptual question. The normative conclusion proposes a change in our language usage – that when we say "natural selection" or "random drift," that we refer to the processes, and that when we want to refer to the outcomes of those processes, we say "the results of

natural selection (random drift)" or "natural selection (random drift) outcomes", or the like. The normative conclusion thus asserts that, as a conceptual matter, natural selection and random drift are distinct, but that as an empirical matter, their outcomes are difficult to distinguish.[9]

I would urge the "normative" conclusion over the "obvious" conclusion, although the less normative among us would clearly prefer the latter. Indeed, I recognize that it is difficult, if not impossible, to change widespread terminological usages. And yet if we do not, I fear that further confusion will be the result.

3.2. An Alternative Solution to the Problem.

As noted above, my position is essentially the same as Hodge's (1987; see also Hodge 1983). However, there are other possible responses to the problem that Beatty raises. For example, Shanahan (1992) argues that conceptually, random drift and natural selection are the ends of a continuum. However, to reach this conclusion, Shanahan must reject heritability as a necessary condition for natural selection. As Endler has argued, this has the effect of trivializing natural selection to the claim that "there are differences among different phenotypes" (Endler 1992, 223). For Rosenberg, the problem is non-existent, since according to Rosenberg, random drift is eliminable from an omniscient account of evolutionary theory (1988), a claim challenged by Millstein (1996). A fifth potential solution[10] is implicit in Brandon and Carson (1996) and Sober (1984). It is to that alternative that I now turn.

Brandon and Carson state: "Conceptually (though often not empirically) drift and natural selection are clearly distinct" (1996, 325, n. 13). This is essentially the conclusion for which I argued above. Brandon and Carson do not offer reasons for this conclusion, but they seem to be relying on a conception of natural selection that is different from Beatty's (or at least different from my interpretation of Beatty[11]). On Beatty's account, which I have largely endorsed (it is his conclusions that I disagree with), both natural selection and random drift are represented as sampling processes, and both are probabilistic.[12] More specifically, the theory of natural selection incorporates probabilities independent of the probabilities that the theory of random drift incorporates. This contrasts with the account offered by Brandon and Carson (1996, 324), in which the formal structure of theory of natural selection is "deterministic" in the sense that Sober elaborates: "*When it acts alone, the future frequencies of traits in a population are logically implied by their starting frequencies and the fitness values of the various genotypes*" (Sober 1984, 110; italics in original). Thus, according to Brandon and Carson, the probabilism of natural selection derives solely from its real-life connection with drift: "*natural selection is indeterministic at the population level because (in real life as opposed to certain formal models) it is inextricably connected with random drift*" (1996, 324; italics in original). The theory of natural selection, on Brandon and Carson's account, is itself not probabilistic, even if in "real life" the phenomenon always is.

Given this account of natural selection, it seems clear why Brandon and Carson reject Beatty's claim that random drift and natural selection cannot be conceptually distinguished. In the case of Beatty's moth example, Brandon and Carson would claim

that the population experienced both natural selection and random drift. They would

argue that to the extent that the change in the population was in accordance with the

expectations of natural selection, the change was due to natural selection, whereas to the

extent that the change deviated from those expectations, the change was due to random

drift. On this kind of view, the concepts of random drift and natural selection are distinct;

natural selection is a process whereby organisms achieve their expected reproductive

success (based on their fitness in a given environment), random drift is a process which

probabilistically deviates the population from those expectations.

   This argument is appealing in that it provides a clean distinction between natural

selection and random drift. It also supports the oft-repeated claim that in most real

populations, both natural selection and random drift are occurring; using Brandon and

Carson's conceptions of random drift and natural selection, natural selection will always

occur with random drift (since no real population is infinite).

   However, the argument makes the same mistake we saw above: it fails to distinguish

processes from outcomes. This failure derives from Brandon and Carson's conceptions

of random drift and natural selection. Brandon and Carson state that drift can be modeled

after the drawing of balls from an urn, in which "each ball in the urn has an equal

probability of being pulled" (1996, 324). With natural selection, on the other hand,

different balls have different chances of being pulled: "some balls are sticky, and some

slippery, with the sticky ones more likely being pulled" (Brandon and Carson 1996, 324).

This is a very nice way of illustrating the difference between random drift and natural

selection as *processes*, with the former being an indiscriminate sampling process and the

latter being a discriminate sampling process. Yet, in Brandon and Carson's example in the same section, random drift and natural selection are spoken of as *outcomes*. Ten balls are chosen from an urn of 10,000 balls, half of which are sticky and half of which are slippery, with the sticky balls twice as likely to be chosen as the slippery. The expected outcome is that 6 2/3 sticky balls and 3 1/3 slippery balls will be chosen, but of course, such an outcome cannot obtain. So, Brandon and Carson ask us to suppose that six sticky balls and four slippery balls are chosen. They then state: "This result [read: outcome] deviates from the expected result and this deviation is drift" (1996, 325). But, regardless of outcome, why should we consider this example to model drift in any way? It does not describe an indiscriminate sampling process; it describes a discriminate sampling process. There is thus an inherent contradiction in Brandon and Carson's conceptions of random drift and natural selection; the "process" conception describes the "deviation" (indeed, the whole example) as natural selection, whereas the "outcome" conception describes the "deviation" as random drift. The proper conclusion to reach, given that the ball drawing is a discriminate sampling process, is that the model describes natural selection alone. Furthermore, natural selection outcomes should be modeled probabilistically. The alternative is to abandon the conceptions of random drift and natural selection as processes, and to define them solely in terms of their outcomes. However, this does not seem to be a very palatable solution, given the usefulness of the discriminate and indiscriminate sampling models.

Furthermore, excluding chance from our conception of natural selection (as Brandon and Carson do) does not seem to fit with longstanding usage of the term in biology.

Darwin, in describing natural selection, asserts: "But if variations useful to any organic being do occur, assuredly individuals thus characterized will have the *best chance* of being preserved in the struggle for life" ([1859] 1964, 127; italics added. See also e.g., pp. 61, 81). Darwin's theory lacked the concept of random drift. Thus, on Darwin's conception of natural selection, chance – in this case the chance that organisms' possessing useful variations may not be successful – is *part* of the concept of natural selection, not separate from it. Ernst Mayr also utilizes a probabilistic conception of natural selection, distinct from the probabilism imparted by random drift:

> Natural selection is a statistical phenomenon; it means merely that the better genotype has a 'better chance of surviving' (Darwin). A light-colored individual in a species of moth with industrial melanism may survive in a sooty area and reproduce, but its chances of doing so are far less than those of a blackish, cryptically colored individual. It happens not infrequently in nature that, for one reason or another, a superior individual fails to reproduce while an inferior one does so abundantly … Natural selection, being a statistical phenomenon, is not deterministic; its effects are not rigorously predictable (1970, 107-8).[13]

In more recent writings, he asserts that "selection and chance are not two mutually exclusive alternatives…there are stochastic perturbations ("chance events") during every stage of the selection process" (1983, 332).

The conception of natural selection inherent in these quotes from Darwin and Mayr is one which many biologists and philosophers of biology would grant, I think: we acknowledge that natural selection is a process where organisms may not be as reproductively successful as one would expect, based upon their traits – they may "by chance" fail to achieve that success. But is random drift this chance element or is the chance element an inherent part of natural selection? Or, to shift the burden of proof to Brandon and Carson, why speak of natural selection as "deterministic"? I think the reason natural selection is often described in this way is to provide a relatively simple model that can be analyzed quantitatively. In the context of population genetics, it is easiest to deal with a model of natural selection where population size is infinite and expected values are realized. Then, when one "complicates" the model by introducing finite population size, random drift suddenly comes into play. Further "complications," such as migration and mutation, can also be added to the model. So the population models serve as step-wise descriptions of evolutionary processes, with increasing complexity. Considering natural selection under the conditions of an infinite population allows biologists to isolate better the role that *other* factors play, such dominant and recessive genes, without complicating factors "getting in the way."

However, our need for simplicity should not override our need for conceptual explanatory power. Natural selection as discriminate sampling reveals processes that occur in the real world, as does random drift as indiscriminate sampling (for example, instances of discriminate and indiscriminate parent sampling and of the bottleneck effect). We can use models that fail to map onto the real world, if we find them useful,

but we should be careful not to confuse them with their more legitimate real-world

cousins.

4. Conclusion.

The neutralist/selectionist debate rages on, with no obvious resolution in sight (see Ridley

1996 for discussion). However, at the very least, it is important to be clear on the terms

of the debate. Beatty's arguments imply that disputants are (at least sometimes) arguing

over nothing, for how can one argue over the prevalence of natural selection and random

drift when the two concepts cannot even be distinguished? Brandon and Carson provide

a coherent way to avoid this unpalatable conclusion, but at the expense of the richness of

the concepts themselves. Natural selection is stripped of its probabalistic character

altogether, while random drift becomes a mathematical deviation from the expectations

of natural selection (rather than seeing it as a process in its own right).

However, the real problem with both of these accounts is that they confuse process

with outcome, and therefore, draw the wrong conclusions. Natural selection and random

drift are sometimes spoken of as processes, and sometimes spoken of as outcomes. This

ambiguity in meaning leads to confusion when we try to distinguish one concept from

another, and it has no doubt created confusion in other areas of evolutionary thinking as

well. But process is not the same as outcome, and so we should be careful to identify

clearly when we are speaking of the process itself, and when we are speaking of its

outcome. Random drift as a process *can* be distinguished from natural selection as a

process, even though random drift as an outcome may not be always be distinguishable from natural selection as an outcome. If we cannot distinguish random drift outcomes from natural selection outcomes, which is more properly an empirical consideration than it is a conceptual consideration, the neutralist/selectionist debate may never be resolved, but at least it will not have been over nothing.

Acknowledgements

Notes

[1] As far as I know, the terms "gamete sampling" and "parent sampling" were first used in Beatty (1984), although he does not make use of this terminology in his 1992 paper.

[2] Throughout this paper, I will speak primarily of "physical differences" rather than "fitness differences" in order to avoid entanglement in the question of how to define fitness, a question that is largely irrelevant to the question at hand.

[3] Of course, the fittest organisms may not have the greatest reproductive success; here we describe only a possible outcome (see discussion below). Similarly, with the process of random drift the outcome may not be that gene frequencies fluctuate from generation to generation; in a large population, gene frequencies may remain relatively constant.

[4] I will question this suggestion below.

[5] The proposition I am asserting (here, and throughout the essay) concerns the structure of the *theory* only. For a discussion of the "probabilistic nature" of the *process* (i.e., for a discussion of the question of whether the evolutionary process is indeterministic) see Rosenberg (1988, 1994), Horan (1994), and Graves, Horan, and Rosenberg (1999), all of whom argue that the evolutionary process is deterministic, and Brandon and Carson (1996), who argue that the evolutionary process is indeterministic. My own position is that given our current state of scientific knowledge, we are currently not in a position to answer this question, and thus, we should remain agnostic (Millstein

2000b). Therefore, the arguments I present in this paper should hold whether the evolutionary process is deterministic or indeterministic.

   [6] Or as being cases of random drift alone, if there are no causally relevant physical differences in the population at all.

   [7] In the finch study, Grant and Grant attribute the outcome largely to natural selection, although their conclusions are by their own admission somewhat less than definitive (Grant and Grant 1989, 233).

   [8] I don't want to rule out the possibility of using probability theory or some other means to specify "consistency" and "repeatability" more precisely, but only to suggest that the problem of distinguishing outcomes is more complex than that of distinguishing processes, and that some outcomes may not be identifiable as natural selection or random drift.

   [9] This issue deserves more discussion than I am able to give it here. For further discussion of the difficulties involved in distinguishing between natural selection and random drift empirically, see Beatty (1987).

   [10]The five potential solutions are those of: 1) Beatty, 2) Hodge/Millstein 3) Shanahan, 4) Rosenberg, and 5) Brandon/Carson/Sober.

   [11] If I am mistaken, and if Beatty's conception of natural selection is the same as Brandon and Carson's, then it is unclear how to make sense of Beatty's claim that it is

difficult to distinguish between the concepts of random drift and natural selection (as will become evident from the discussion below).

[12] Again, I am not arguing for the indeterminism of the process of natural selection; in this section I am making only a much weaker claim concerning the probabilistic structure of the theory (see note 5).

[13] Mayr is clearly not considering random drift here; he brings random drift and other forms of drift into his account at a later point in the chapter (1970, 120-8).

References

Beatty, J.: 1984, 'Chance and Natural Selection', *Philosophy of Science* 51, 183–211.

Beatty, J.: 1987, 'Natural Selection and the Null Hypothesis', in J. Dupré (ed.), *The Latest on the Best: Essays on Evolution and Optimality*, MIT Press, Cambridge, MA, pp. 53–75.

Beatty, J.: 1992, 'Random Drift', in E. F. Keller and E. A. Lloyd (eds.), *Keywords in Evolutionary Biology*, Harvard University Press, Cambridge, MA, pp. 273–281.

Brandon, R.: 1990, *Adaptation and Environment*, Princeton University Press, Princeton, NJ.

Brandon, R. and S. Carson: 1996, 'The Indeterministic Character of Evolutionary Theory: No 'No Hidden Variables' Proof But No Room for Determinism Either', *Philosophy of Science* 63, 315–337.

Darwin, C.: [1859] 1964, *On the Origin of Species: A Facsimile of the First Edition*, Harvard University Press, Cambridge, Massachusetts.

Dodson, E. O. and P. Dodson: 1985, *Evolution: Process and Product*, Prindle, Weber and Schmit, Boston.

Endler, J. A.: 1986, *Natural Selection in the Wild*, Princeton University Press, Princeton, NJ.

Futuyma, D. J.: 1988, 'Sturm Und Drang and the Evolutionary Synthesis', *Evolution* 42, 217–226.

Futuyma, D. J.: 1986, *Evolutionary Biology*, Sinauer Associates, Sunderland.

Grant, B. R. and P. R. Grant: 1989, *Evolutionary Dynamics of a Natural Population: The Large Cactus Finch of the Galápagos*, University of Chicago Press, Chicago.

Graves, L., B. L. Horan, et al.: 1999, 'Is Indeterminism the Source of the Statistical Character of Evolutionary Theory?', *Philosophy of Science* 66, 140–157.

Hodge, M. J. S.: 1983, 'The Development of Darwin's General Biological Theorizing', in D. S. Bendall, (ed.), *Evolution from Molecules to Men*, Cambridge University Press, Cambridge, pp. 43-62.

Hodge, M. J. S.: 1987, 'Natural Selection as a Causal, Empirical, and Probabilistic Theory', in L. Krüger (ed.), *The Probabilistic Revolution*, MIT Press, Cambridge, MA, pp. 233–270.

Horan, B. L.: 1994, 'The Statistical Character of Evolutionary Theory', *Philosophy of Science* 61, 76–95.

Mayr, E.: 1970, *Populations, Species, and Evolution*, The Belknap Press of Harvard University Press, Cambridge, Massachusetts.

Mayr, E.: 1983, 'How to Carry Out the Adaptationist Program?', *American Naturalist* 121, 324–334.

Millstein, R. L.: 1996, 'Random Drift and the Omniscient Viewpoint', *Philosophy of Science* 63: S10-S18.

Millstein, R. L.: 2000a, 'Chance and Macroevolution', *Philosophy of Science* 67, 603-624.

Millstein, R. L.: 2000b (unpublished), 'Is the Evolutionary Process Deterministic or Indeterministic? An Argument for Agnosticism', Presented at the Philosophy of Science Association conference, November 2000 in Vancouver, Canada.

Ridley, M.: 1996, *Evolution, 2nd. Edition*, Blackwell, Cambridge, MA.

Rosenberg, A.: 1988, 'Is the Theory of Natural Selection a Statistical Theory?', *Canadian Journal of Philosophy (Suppl).* 14, 187–207.

Rosenberg, A.: 1994, *Instrumental Biology or the Disunity of Science*, University of Chicago Press, Chicago.

Shanahan, T.: 1992, 'Selection, Drift, and the Aims of Evolutionary Theory', in P. Griffiths (ed.), *Trees of Life: Essays in Philosophy of Biology*, Kluwer, Dordrecht, pp. 131–161.

Sober, E.: 1984, *The Nature of Selection*, MIT Press, Cambridge, MA.