The Nature of Evolutionary Biology: At the Borderlands Between Historical and Experimental Science

Massimo Pigliucci

> Even today a good many distinguished minds seem unable to accept or even to understand that from a source of noise natural selection could quite unaided have drawn all the music of the biosphere.
>
> <div align="right">Jacques Monod</div>

1 On the Scientific Status of Evolutionary Theory

The scientific status of evolutionary theory seems to be more or less perennially under question. I am not referring here (just) to the silliness of young Earth creationism (Pigliucci 2002; Boudry and Braeckman 2010), or even of the barely more intellectually sophisticated so-called Intelligent Design theory (Recker 2010; Brigandt this volume), but rather to discussions among scientists and philosophers of science concerning the epistemic status of evolutionary theory (Sober 2010). As we shall see in what follows, this debate has a long history, dating all the way back to Darwin, and it is in great part rooted in the fundamental dichotomy between what French biologist and Nobel laureate Jacques Monod (1971) called chance and necessity—i.e., the inevitable and inextricable interplay of deterministic and stochastic mechanisms operating during the course of evolution. In turn, this discussion reaches as far as our very concept of what counts as a science and why, with the perennial struggle between primarily or exclusively experimental sciences (e.g., physics) on one hand and chiefly historical sciences (e.g., paleontology) on the other. While the two issues (deterministic vs. stochastic phenomena and experimental vs. historical sciences) are

M. Pigliucci (✉)
Philosophy Program, The Graduate Center, City University of New York,
New York, NY, USA
e-mail: massimo@platofootnote.org

logically distinct, as we shall see they overlap in practice, and for philosophically interesting reasons.

Hull (1973) pointed out that evolutionary theory got under the fire of philosophers of science from the onset, with the publication of Darwin's (1859) masterpiece. In particular, Darwin's famous "long argument" laying out the foundations of the new field of evolutionary biology, was seen as ill-fitting with both of the major philosophical views of how science works that were being debated in late nineteenth century Victorian England. This is known as the great induction debate, and featured a who's who of early philosophy of science, with John Stuart Mill and William Whewell (the man who coined the term "scientist" in 1834), both attempting to improve on the model of scientific reasoning articulated by their predecessor, Francis Bacon. Famously, Bacon (1620/2000) had published a provocatively titled book, *The New Organon*, in which he set out to criticize what he took as Aristotle's view (expressed in the latter's *Organon*) that science proceeds by way of deductive reasoning. Bacon thought that deduction wouldn't do, because the premises of any deductive argument have to be arrived at by way of empirical evidence, which means that at bottom science relies on a different type of reasoning, induction.

The problem is that there are several types of induction, and the one Bacon mostly focused on, enumerative induction, suffers from severe problems of its own. Essentially, enumerative induction is the everyday practice of generalizing from a small number of instances to a broader context, a practice that Hume (1748/1952) quickly realized is without independent logical foundations and open to spectacular failures. Hence, the debate between Mill and Whewell on how to improve on Bacon's proposal. Mill thought that enumerative induction could be strengthened by invoking a law of universal causation, according to which similar effects are produced by similar causes, an example of parsimony applied to scientific explanations. Mill also added eliminative induction, sometimes referred to as "strong inference" (Platt 1964), where alternative hypotheses are eliminated in succession, based on stringent tests or crucial experiments.

Both Bacon and Mill gave preeminence to data over hypotheses, while Whewell's approach emphasized the idea that a scientific investigation has to start with some hypothesis, even if approximate, to serve as a heuristic guide to data collection and further hypothesizing. Whewell thought that progress was made by consilience, a situation where the data consistently point to one particular hypothesis being true over its rivals. This way of reasoning is also known with the somewhat strange name of abduction, or inference to the best explanation.

Darwin got in trouble with both Mill and Whewell because they both regarded Darwin's work as an example of deductive reasoning, not of induction (of any kind). According to Mill and Whewell's understanding, what Darwin had done was to arrive at a hasty generalization based on a small number of observations, proceeding then (deductively) to derive consequences from them, and finally collecting data for decades to back up his hasty conclusions. This superficially appeared to be almost a textbook example of what Bacon said one should not do in science. Evolutionary biology, particularly when compared to the already

successful Galilean and Newtonian physics, had already gotten off on the wrong foot. Needless to say, Darwin was taken aback by all this, particularly by Whewell's criticism. Indeed, in apparent frustration, Darwin wrote to a friend: "How odd it is that anyone should not see that all observation must be for or against some view if it is to be of service!" (Darwin, F. 1903, p. 240), a sentiment that is actually perfectly consonant with Whewell's idea of inference to the best explanation (and, indeed, is also compatible with Mill's eliminative induction, to a point). In fact, Michael Ruse (1975, 2000) has persuasively argued that Darwin consciously tried to develop his theory in accordance with the best philosophy of science of his time, particularly following the views of Whewell and John Herschel.

The reason these early skirmishes are important is because they set the stage for much discussion over the following century and a half about the goals and methods of evolutionary biology. Moreover, as we shall see below, the distinction between eliminative induction and inference to the best explanation parallels, roughly, the distinction between the methods of the experimental and historical sciences as determined by their different epistemologies.

2 The Fisher-Wright Debates and the Importance of Stochastic Events in Evolution

The next big battle for the soul of evolutionary biology was the famous, decade-long, debate between two of the founding fathers of population genetics, Ronald Fisher and Sewall Wright. The history of this debate, and the role it played in population genetics, is both complex and fascinating, and it has been examined in detail elsewhere (Provine 2001). Fisher was convinced of the absolute preeminence of natural selection in shaping organismal evolution, so much so that he consciously modeled his famous fundamental theorem (Frank and Slatkin 1992; Okasha 2008) after the second principle of thermodynamics, one of the most successful laws formulated within the dominant experimental science, physics.

Wright, on the other hand, was attracted by the complexity and messiness of biology, and his emphasis was always on non-linear, non-additive genetic effects (epistasis, pleiotropy) as well as, most famously, on the role of genetic drift in evolution. Wright built drift into his innovative theory of "shifting balance", meant to account for how deterministic selection and stochastic drift combine to allow populations to explore new "peaks" in what Wright defined as the "adaptive landscape" (Pigliucci 2008). The status of shifting balance in evolutionary theory is itself under perennial discussion (Wade 1992; Coyne et al. 1997), though there seems to be an emerging sense that the mechanism, while theoretically feasible, is unlikely to play a major role in the evolution of actual biological lineages. Still, the concept of drift has become incorporated into standard population genetic theory, and has even been the focus of detailed, and still unsettled, philosophical analyses (e.g., Pigliucci and Kaplan 2006; Millstein et al. 2009).

Indeed, in a recent attempt at a major shift in perspective, Lynch (2007) has forcefully pushed the idea that a great deal of genome evolution at the structural level (e.g., how gene networks are formed and change over time) is likely the result of drift, not of selective mechanisms, as usually assumed. Whether and to what extent Lynch is correct remains to be seen, but the "problem" of distinguishing drift from selection in now standard in biological research (e.g., Chapuis et al. 2008; Hofer et al. 2008).

It is important to understand that "drift," although often portrayed as a "force" affecting the equilibrium state of natural populations (together with the other four canonical "forces" that can cause displacement from Hardy-Weinberg equilibrium, selection, mutation, migration and recombination: Hartl and Clark 1997, pp. 48–52; see also Depew this volume), actually is a beast of a very different kind from natural selection. Perhaps the best way to make the point clear is to use the standard analogy between genetic drift and the kind of statistical drift one commonly observes when examining series of coin tosses of equal length. Assuming that the coins are fair (i.e., there is no weight biasing the outcome toward either tail or head), the expected outcome of a series is, of course, 50–50 %. This, however, is true only asymptotically, as the series reaches infinite length—a direct analogy with the oft-made assumption of "infinite" (i.e., not affected by drift) population size in population genetic models. The shorter the series, the more likely it is that the actual outcome is going to deviate (in either direction, with random frequency) from the expected one. The likelihood and intensity of these deviations become increasingly high as the coin toss series becomes shorter and shorter. It is in a very similar sense that biologists speak of drift getting "stronger" in smaller and smaller populations.

But notice a rather odd thing about this account of drift: if one focuses on individual coin tosses, no quantity is changing at all. For each individual toss the chances of landing tail or head is always the same (*ceteris paribus* as far as environmental conditions are concerned, including the strength and mode of the toss). There is no mystery here, of course: drift is a phenomenon that manifests itself only when a *population* of objects is concerned; it is undefined at the individual level. But this is different from all the other "forces" that can push a population off Hardy-Weinberg equilibrium: mutations, recombination, migration and selection happen at the *individual* level, though of course their *outcomes* may be conveniently measured statistically for entire populations. There is, therefore, something strange going on when biologists want to know the relative strength of selection vs. drift in a population, considering that they are comparing mechanisms acting at two different levels of the biological hierarchy.

Jonathan Kaplan and I (Pigliucci and Kaplan 2006) have suggested a way to conceptualize what is happening here, an approach that might be helpful as part of a general understanding of the roles of chance and necessity in evolution. Instead of thinking of drift as a force antagonistic to selection, we can conceptualize it as a measure of the "error" surrounding the expected evolutionary change caused by selection. So, if the target of selection is a particular phenotypic value of a particular trait (in a particular environment, of course), then there is a probability distribution that tells us how likely the population actually is to land on that phenotypic target.

The smaller the population (the higher the drift), the broader the error will be, with increasing chance that the population will end up anywhere in an expanding circle centered on the actual target. This is a different, and we think more satisfying, account of how selection and drift interact, because it gets away from the "force" metaphor (another reason for doing which is that drift does not have the properties of a vector, so it cannot reasonably be represented as a force; see also Depew this volume about whether natural selection should be described as a force). It also makes explicit that we can answer the drift vs. selection question only in terms of *outcomes* at the population level, the question being meaningless at the individual level because individuals experience selection, but not drift.

3   Gould and the Project for a Nomothetic Evolutionary Biology

The latter part of the twentieth century saw the opening of a different front in the ongoing discussion about the relative role of chance events in evolutionary biology, as well as about the status of the discipline as historical and yet scientifically fully mature. The main charge was led by Stephen Gould and his associates, with a series of papers that unleashed decades of debates and new research—a voluminous output that makes for what philosopher Imre Lakatos called a successful (as opposed to a degenerate) research program in the light of his treatment of the nature of science (Lakatos 1978).

   The opening salvo by Gould and company was the famous 1972 paper on punctuated equilibria (Eldredge and Gould 1972), where the standard Darwinian view of gradual evolution was challenged and, by implication, the role of natural selection in shaping long-term evolution somewhat curtailed. Gould then developed that theme in his highly influential book, *Ontogeny and Phylogeny* (Gould 1977), where the idea is put forth that developmental constraints play a major role in both antagonizing and sometimes even facilitating natural selection. Because constraints are themselves the result of past history, and hence also of stochastic events, this helped shape an alternative and broader view of evolutionary phenomena, one where organismal biology (including of course paleontology) would reclaim a place at the high table, so to speak (Prothero 2009) in virtue of the renewed focus on historicity and chance events. Indeed, Gould even partially rethought the role of natural selection itself by introducing a new term in the evolutionary jargon, exaptation, to indicate situations where currently advantageous traits result from selection co-opting previously existing structures, which themselves evolved either for other functions or as the result of constraints (Gould and Vrba 1982; see also Forber this volume). In a sense, exaptation is what one gets when chance and necessity work sequentially.

   Gould's most conscious attempt at articulating a new view for paleontology in particular, and for historical biological science in general, came with his paper on nomothesis (Gould 1980). In philosophy of science nomothetic means law-like (see Lange this volume), which is somewhat puzzling because Gould's emphasis

throughout his career was on the importance of stochastic, not deterministic causes in evolution. But Gould was also aware that one needs to combine stochasticity and determinism if one wishes to obtain a reasonably complete view of biological phenomena. His attempt at nomothesis took the shape of a novel approach to the study of why certain forms and not others appear throughout the history of any given lineage. The basic observation is obvious enough: not all conceivable forms of, say, shelled animals are in fact found in the fossil record. Why not? There is more than one conceivable answer to the question. Perhaps some forms did evolve, briefly, but went extinct. Or, perhaps, some shell shapes may simply be intrinsically disadvantageous, regardless of the specific environment, and therefore are constantly selected against. Finally, forms would be selected in favor if they could be generated, but perhaps the genetic-developmental systems, characteristic of certain lineages, simply are incapable of producing the right kind of variation.

The idea began with paleontologist David Raup (Raup and Gould 1974), who formulated a theoretical space of shell forms (whimsically called by Gould "the cube of life") and then filled it in with actual living or fossil shells to see which parts of the space have historically been occupied and how frequently. What is important to notice here is that Raup and Gould did not construct the space of forms *a posteriori*, i.e. from actual observations of shells. It was derived a priori using a simple equation that describes the growth curve of a shell. The equation has three basic parameters, which constitute the three axes of the cube of life, and it is the variation of these parameters that generates all hypothetical shell forms. Only afterward the researcher plots actually observed shell types on the same space, and it is the comparison and differences between the hypothetical space and the real organisms that is of interest. Of particular relevance, of course, are the areas of the cube of life occupied by theoretically possible, but never realized, shell shapes. Once these are identified, the scientist can bring in the result of functional analyses (for instance about buoyancy, or strength of the shell in response to predators, etc.) to account for at least some of the mismatches between hypothetical and actual spaces.

A particularly interesting example of the application of this approach is recounted in detail by McGhee (2006). The work was again started by Raup (1967), who explored a mathematical-geometrical space of ammonoid forms defined by two variables: W, the rate of expansion of the whorl of the shell; and D, the distance between the aperture of the shell and the coiling axis. As McGhee shows in his detailed discussion of this example, Raup arrived at two simple equations that can be used to generate pretty much any shell morphology that could potentially count as "ammonoid-like," including shells that—as far as we know—have never actually evolved in any ammonoid lineage. Raup then moved from theory to empirical data by plotting the frequency distribution of 405 actual ammonoid species in W/D space and immediately discovered two interesting things: first, the distribution had an obvious peak around $0.3 < D < 0.4$ and $W \approx 2$. Consider that this kind of peak is not a direct measure of fitness or adaptation, it is simply a reflection of the actual occurrence of certain forms rather than others. Second, the entire distribution of ammonoid forms was bounded by the $W = 1/D$ hyperbola, meaning that few if any species crossed that boundary on the morphospace. This was interesting, since the 1/D line represents the limit in morphospace

where whorls still overlap with one another. This means that for some reason very few ammonites ever evolved shells in which the whorls did not touch or overlap.

Raup's initial findings were intriguing, but they were missing a sustained functional analysis that would account for the actual distribution of forms in W/D space. Why one peak, and why located around those particular coordinates? Here is where things become interesting and the morphospace approach delivers much more than just heuristic value. John Chamberlain, a student of Raup, carried out experimental work to estimate the drag coefficient of the different types of ammonoid shells. His first results (Chamberlain 1981) clarified why most actual species of ammonoids are found below the $W = 1/D$ hyperbola: as one would expect, shells with overlapping whorls have a significantly lower drag coefficient, resulting of course in more efficiently swimming animals.

However, Chamberlain also found something else rather intriguing: the experimental data suggested that there should be *two* regions of the W/D morphospace corresponding to shells with maximum swimming efficiency, while Raup's original frequency morphospace detected only one peak. It seemed that for some reason natural selection evolved one peak, but not the other. Four decades had to pass from Raup's initial paper for the mystery of the second peak to be cleared up: the addition of 597 new species of ammonoids to the original database showed that indeed the second peak had also been occupied! This is a rather spectacular case of confirmed prediction in evolutionary biology, not exactly a common occurrence, particularly in paleontology, and a superb vindication of Raup's and Gould's idea that historical disciplines where stochastic events play a major role can still be nomothetic in an interesting and sometimes even decidedly surprising way.

4   The Modern Study of Chance vs. Necessity

Part of the problem with the antithesis between chance and necessity—to put it as Monod did—or stochasticity and determinism, is that the best we can usually do is to study their relative importance only retroactively, attempting to determine after the fact the respective contributions of selection, drift and constraints on already realized evolutionary trajectories. During their long sparring careers, Richard Dawkins and Stephen Gould have often argued about what would happen if one could "rewind the tape of life," to use Gould's metaphor (Beatty 2006; Sterelny 2007). The outcome of such a thought experiment was very different according to the two antagonists: Dawkins suggested that we, as *Homo sapiens* might not end up evolving again, but *surely* some sort of hyper-intelligent, big-brained biped would. Gould, on the contrary, suggested that for all we know not even vertebrates might have re-evolved, depending on minor vagaries of the Burgess Shale fauna, back in the Middle Cambrian (540 million years ago). Or perhaps planet Earth may never have gotten past a biosphere made entirely of bacteria and blue algae. Of course, the problem is that there actually is no way to rewind the tape of life, so that any such discussion amounts to speculation largely reflecting the intellectual prejudices of the speculators.

However, recent research on long-term laboratory evolution in microorganisms has offered us the possibility to replay short segments of the tape of life, over and over again while manipulating the conditions to see what happens. A full review of that literature is beyond the scope of this chapter, but I will briefly discuss a couple of examples, to provide a flavor of the possibilities opened up by this approach—arguably a powerful window into the relative importance of randomness and determinism, at least as far as simple and rapidly evolving organisms are concerned.

The first example comes from work done on *Chlamydomonas reinhardtii*, a single celled green alga, by Graham Bell's group at McGill University (Collins et al. 2006). The group began with populations of *Chlamydomonas* that had previously been selected for survival at high levels of CO_2, imposing on them new selective pressures to adapt them either to current or to Pleistocene CO_2 levels (i.e., lower and much lower than the starting point levels, respectively). Bell and colleagues consciously patterned their experiment after Gould's "tape of life" metaphor, even using the phrase in the title of the paper. Their results showed that *Chlamydomonas* had no trouble re-adapting to current CO_2 levels, or indeed to Pleistocene levels. In that sense, the tape of life could indeed be rewound with similar results. However, the process also produced a range of phenotypes in the replicate populations, an outcome that Collins et al. interpreted as indicative of the importance of stochastic events in shaping the details of the evolutionary trajectories. As they put it: "There was no effect of selective history on adaptation […] phenotypes were evolutionarily reversible. [However] Adaptation produced a range of phenotypes, suggesting that chance rather than selective history contributes to the divergence of replicate populations in this system" (Collins et al. 2006, p. 1392).

The approach taken by Woods et al. (2006), within the context of work done by Richard Lenski's lab at Michigan State University, was different yet aiming at the same problem of replicability of evolutionary outcomes. Lenski and colleagues began with 12 identical lines of the bacterium *Escherichia coli*, and let them evolve under the same conditions for about 20,000 generations. The idea was to track the molecular evolution of four pre-identified gene sequences to see if and how they would diverge in the replicated lines. Woods et al. found that two of the genes had substitutions in all populations, while the other two genes had substitutions in several but not all populations. Statistical tests were congruent with the conclusion that the parallel changes in the candidate genes were driven by natural selection, but the researchers found few cases in which the exact same mutations were substituted—thereby documenting a significant role also for stochastic events.

Both experiments were of course likely to underestimate the relevance of stochasticity, because they were carried out by design under uniform environmental conditions. Yet, even under those conditions chance clearly affected the evolutionary outcome. It also has to be noted that one of Gould's major sources of interference with natural selection, developmental constraints, could not play any role in either experiment, for the simple reason that neither *Chlamydomonas* nor *E. coli* have any development to speak of. Regardless, Woods et al. found differences among the replicate populations that were not found in a conceptually similar experiment in viruses, presumably because the *genetic* constraints on viral sequences are stronger

than those acting on the much larger genome of *E. coli*. This is consistent with the general idea that larger genomes (and eventually actual developmental systems) allow more room for stochastic events, as well as for a more complex interplay of chance and necessity. At the same time, though, it is hard to imagine how similar experiments could be done on multicellular eukaryotes, the sort of organisms that are really at the center of Dawkins-Gould type debates. Nonetheless, these and more recent (e.g., Saxer et al. 2010) experiments clearly provide us with an empirical, if limited, handle on the question.

5   The Philosophical Context: Cleland's Analysis

Much of the preceding discussion—like any discussion on the nature of science—reflects a certain philosophy of what it means to be doing experimental or historical science. It is time to explicitly deal with that philosophy. An excellent framework for it is provided by a paper published by Carol Cleland (2002; see also Cleland 2011 for her recent elaboration and rebuttal to critics) on the epistemic differences between historical and experimental sciences. Cleland's pivotal idea is that the two types of science are separated by what she calls an asymmetry of overdetermination. Building on previous work by David Lewis, she explains that "the basic idea is that localized present events overdetermine their causes and underdetermine their effects." She elucidates the concept by considering the example of a crime being investigated. Once committed, a crime leaves a number of historical traces, no matter how careful the perpetrator was in erasing as many of them as possible. All it takes for a criminal investigator to figure out what happened is a relatively small number of traces that clearly enough point toward a particular sequence of events. The investigator would then be using a type of induction known as inference to the best explanation to pinpoint the culprit (the same one we have seen argued for by Whewell early on). Conversely, the simple act of *not* committing the crime obviously instantly erases the possibility of any historical trace to be left around. Few currently available clues *overdetermine* a past event, while so many futures are possible given a particular current state of things that the latter *underdetermines* the range of futures.

Cleland cashes in this asymmetry of overdetermination by arguing that—contrary to popular wisdom (and to the opinion of many practicing scientists)—there is nothing inherently epistemically superior about experimental over historical science. This is because of two consequences of overdetermination. On the one hand, while experimental scientists have the ability to strictly control the conditions of their experiments, it is that very strictness that limits the scope of applicability of their results: as soon as one widens the settings of a given experiment, different factors begin to interact with each other in complex ways, quickly leading to a large number of possible future outcomes; in other words, predictability is purchased at the expense of generality. On the other hand, while historical traces constantly decay through time, and may disappear forever, the historical scientist often needs only

a small amount of them to arrive at a sufficiently accurate reconstruction of what happens—just like the criminal investigator in the hypothetical example of the impossibility of a perfect crime.

To make things more concrete, Cleland's account makes sense of some surprising limitations of experimental science, as well as some spectacular successes of historical science. In the first case, it is notable, for instance, that non-equilibrium thermodynamics (for example in its applications to atmospheric physics and climate science) quickly reaches a limit in terms of predictive ability, where complex mathematical models are incapable of generating more than very approximate statistical predictions about the future behavior of complex systems, predictions often accompanied by rapidly expanding margins of error. In the second case, however, we have for example the success of paleontologists in determining that an extraterrestrial body of massive proportions hit the Earth 65 million years ago, contributing to the extinction of countless numbers of species, chief among them the dinosaurs. The impact was suspected once geologists discovered a worldwide thin layer of iridium in rocks datable to the K/T (Cretaceous/Tertiary) boundary. This led to a search for a crater, the remnants of which were eventually identified off the Yucatan peninsula via satellite imagery. From there, geologists could calculate the size and direction of the impact, and therefore make fresh predictions concerning additional historical traces, for instance those left by the tsunamis that must have hit the western coast of Mexico as a result of the asteroid crush. Sure enough, those traces were found, leading to even more confidence in the conclusion that "the crime" had indeed taken place in the way it had been hypothesized.

Cleland's framework is particularly helpful for a re-evaluation of the chance vs. necessity issue in evolutionary biology. By its very nature, evolutionary studies are *both* experimental (consider again the examples from the Bell and Lenski labs discussed above), as well as historical (Gould's and Raup's reconstruction of morphospaces affecting the evolution of shelled animals). While biologists may not be able to conduct meaningful experiments on *the same* systems for which they have abundant historical information (*E. coli* are certainly different from ammonoids), the fact that they do have access to vast repositories of historical information *and* that they can conduct controlled experiments, albeit within limits, puts evolutionary biology in a position of epistemic advantage over purely historical and possibly even purely experimental sciences, as long as its dual historical/experimental nature is properly understood and adequately exploited by its practitioners.

6   Conclusion: Chance and Necessity Within the Extended Synthesis

Jacques Monod's (1971) analysis of evolutionary theory and the philosophy of biology—with his dualism of chance and necessity—appeared at a strange juncture in the history of biology and of philosophy of science. In the late 1960s and early

1970s evolutionary biologists were satisfied with the Modern Synthesis of the 1930s and 1940s (Huxley 1942/2010), and both the explosion of evo-devo (evolution of development: Love 2009; Love this volume) and of evolutionary genomics (Lynch 2007) were still to come. At the same time, philosophy of science was largely dominated by the philosophy of physics, with philosophy of biology emerging as a mature field only in the 1980s and 1990s.

The situation has, of course, changed dramatically since. Not only is philosophy of biology arguably the dominant sub-field in philosophy of science, or at the very least a major player, but evolutionary biology finds itself again facing a moment of re-evaluation of the basic structure of its theory. There are no serious suggestions that the original Darwinian core, or its expanded outlook within the Modern Synthesis, are in danger of being undermined or rejected, *pace* the creationists. However, an increasing number of biologists and philosophers are convinced that a new, Extended Synthesis is taking place in evolutionary biology (Pigliucci and Müller 2010).

It remains to be seen what sort of Extended Synthesis will emerge over the next decade or so, but one of the elements that seems pretty much certain to be a part of it is in fact a serious and more nuanced consideration of the balance between deterministic and stochastic events in evolution. This is arguably in good part a result of Gould's legacy (particularly of his still somewhat under-appreciated last book: Gould 2002), but also of the onset of evo-devo, which puts constraints at the forefront of its agenda (Müller 2007), and of increasing evidence that large aspects of genomic evolution are more likely the result of stochastic events than of fine-tuning orchestrated by natural selection (Lynch 2007).

Gould (2002) in particular argued that we can follow the evolution of evolutionary theory by considering how our ideas about natural selection—one of the two cardinal concepts that make up the original Darwinian core (the other one being common descent)—change in light of three criteria: agency, efficacy and scope. *Agency* refers to where natural selection acts: genes, individuals, groups, species, etc. In this respect, the moves from core Darwinism to the Modern Synthesis and to an Extended Synthesis have seen the agency of selection increase, as more levels of biological organization are recognized as being possible targets of selection (Okasha 2006). *Efficacy* refers to the relative power of natural selection when compared to other evolutionary mechanisms. Here the story is different, since the Modern Synthesis introduced several additional evolutionary mechanisms, reflected in the core mathematical treatment of population genetics (Hartl and Clark 1997). Particularly due to the work of Sewall Wright and his followers, it is clear that drift is here to stay as a major counter to the efficacy of selection. Gould's emphasis on developmental constraints falls in the same category, the two efforts resulting in an augmented role of chance in evolution. Finally, s*cope* is the degree to which microevolutionary phenomena can be extrapolated to macroevolutionary ones. Notoriously, we saw no change here when biologists moved from Darwinism to the Modern Synthesis, particularly after the so-called "hardening" of the Synthesis that involved a much reduced role of paleontology. But the Extended Synthesis is bound

to take seriously the work of Gould and his followers (Jablonski 2005), resulting in a partial decoupling of micro- and macro-evolution, a decoupling realized by the intervention of stochastic processes (which make it impossible to simply extrapolate macroevolutionary patters from microevolutionary phenomena), and hence by a surprisingly anti-reductionist role of stochasticity (after all, asteroids wiping out entire ecosystems are hardly the stuff of molecular biology).

These are exciting times for evolutionary biology, both in terms of empirical discoveries (evo-devo, comparative genomics) and conceptual advances (e.g., discussions of evolvability, emergent complexity, and the like: Brigandt 2007). And while it is true, as Monod said, that "even today a good many distinguished minds seem unable to accept or even to understand that from a source of noise natural selection could quite unaided have drawn all the music of the biosphere," that understanding is getting richer and deeper, and it still hinges on taking seriously the dichotomy and interaction between randomness and determinism.

What, then, are the implications of all of this for the teaching of evolutionary biology? The standard approach, at both pre-college and introductory college levels, is to teach students largely about micro-evolution, probably because of the inherent prestige of population genetics—the most mathematical of the sub-branches of evolutionary biology. Students are therefore exposed to the centrality of natural selection as an evolutionary mechanism, as well as to the focal idea that evolution can be defined in terms of changes in gene frequencies.

What does not get much attention is the broader picture of evolution offered by a consideration of macro-evolutionary patterns and processes, and particularly an explicit emphasis on stochasticity and on the consequences of rare but high impact events that take place during the history of life. Indeed, Catley (2006) has identified this as a major deficiency in the teaching of evolution, and has proposed a radical shift toward a much more balanced exposure of students to both micro- and macro-evolutionary concepts. The sort of historical and philosophical discussions about chance vs. necessity that I have briefly outlined in this chapter provide precisely the sort of additional intellectual exposure that Catley is advocating.

Moreover, the standard approach to teaching biology is notoriously missing in conceptual and philosophical themes of the type developed, for instance, through Cleland's analysis of the asymmetry of overdetermination. This is unfortunate not just from the standpoint of general educational principles (e.g., that science education—especially for students who are not likely to pursue science at a more advanced level—should be about the "big picture" and not a host of technical details). It is a deficiency highlighted by empirical evidence that a better appreciation of the nature of science itself is correlated with and facilitates the development of better understanding of evolutionary biology (Nadelson and Sinatra 2010). It may very well be, as Dobzhansky famously put it, that nothing (in the teaching of) biology makes sense except in the light of evolution. But it makes increasingly less sense to teach evolution within a narrow perspective and without proper historical and philosophical contexts.

References

Bacon, F. 1620/2000. *The new organon*. Cambridge: Cambridge University Press.

Beatty, J. 2006. Replaying life's tape. *Journal of Philosophy* 103: 336–362.

Boudry, M., and J. Braeckman. 2010. Immunizing strategies and epistemic defense mechanisms. *Philosophia* 39(1): 145–161.

Brigandt, I. 2007. Typology now: Homology and developmental constraints explain evolvability. *Biology and Philosophy* 22: 709–725.

Catley, K.M. 2006. Darwin's missing link—A novel paradigm for evolution education. *Science & Education* 90: 767–783.

Chamberlain, J.A. 1981. Hydromechanical design of fossil cephalopods. *Systematics Association Special Volume* 18: 289–336.

Chapuis, E., G. Martin, and J. Goudet. 2008. Effects of selection and drift on G matrix evolution in a heterogeneous environment: A multivariate Qst–Fst test with the freshwater snail *Galba truncatula*. *Genetics* 180: 2151–2161.

Cleland, C.E. 2002. Methodological and epistemic differences between historical science and experimental science. *Philosophy of Science* 69: 474–496.

Cleland, C.E. 2011. Prediction and explanation in historical natural science. *The British Journal for the Philosophy of Science* 62: 551–582.

Collins, S., D. Sultemeyer, and G. Bell. 2006. Rewinding the tape: Selection of algae adapted to high CO_2 and current and Pleistocene levels of CO_2. *Evolution* 60: 1392–1401.

Coyne, J.A., N.H. Barton, and M. Turelli. 1997. A critique of Sewall Wright's shifting balance theory of evolution. *Evolution* 51: 643–671.

Darwin, C. 1859. *On the origin of species by means of natural selection: Or, the preservation of favoured races in the struggle for life*. London: John Murray.

Darwin, F. 1903. *More letters of Charles Darwin. A record of his work in a series of hitherto unpublished letters*. London: John Murray.

Eldredge, N., and S.J. Gould. 1972. Punctuated equilibria: An alternative to phyletic gradualism. In *Models in paleobiology*, ed. T.J.M. Schopf, 82–115. San Francisco: Freeman.

Frank, S.A., and M. Slatkin. 1992. Fisher's fundamental theorem of natural selection. *Trends in Ecology & Evolution* 7: 92–95.

Gould, S.J. 1977. *Ontogeny and phylogeny*. Cambridge: Harvard University Press.

Gould, S.J. 1980. The promise of paleobiology as a nomothetic, evolutionary discipline. *Paleobiology* 6: 96–118.

Gould, S.J. 2002. *The structure of evolutionary theory*. Cambridge: Harvard Press.

Gould, S.J., and E.S. Vrba. 1982. Exaptation—A missing term in the science of form. *Paleobiology* 8: 4–15.

Hartl, D.L., and A.G. Clark. 1997. *Principles of population genetics*. Sunderland: Sinauer.

Hofer, T., N. Ray, D. Wegmann, and L. Excoffier. 2008. Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Annals of Human Genetics* 73: 95–108.

Hull, D. 1973. *Darwin and his critics: The reception of Darwin's theory of evolution by the scientific community*. Cambridge: Harvard University Press.

Hume, D. 1748/1952. *An enquiry concerning human understanding*. Charleston, SC: Forgotten Books.

Huxley, J. 1942/2010. *Evolution: The modern synthesis: The definitive edition*. Cambridge: MIT Press.

Jablonski, D. 2005. Evolutionary innovations in the fossil record: The interaction of ecology, development and macroevolution. *The Journal of Experimental Zoology* 304B: 504–519.

Lakatos, I. 1978. *The methodology of scientific research programmes*. Cambridge: Cambridge University Press.

Love, A. 2009. Marine invertebrates, model organisms, and the modern synthesis: Epistemic values, evo-devo, and exclusion. *Theory in Bioscience* 128: 19–42.

Lynch, M. 2007. *The origins of genome architecture*. Sunderland: Sinauer.

McGhee, G.R. 2006. *The geometry of evolution: Adaptive landscapes and theoretical morphospaces*. Cambridge: Cambridge University Press.

Millstein, R.L., R.A. Skipper, Jr., and M.A. Dietrich. 2009. (Mis)interpreting mathematical models: Drift as a physical process. *Philosophy & Theory in Biology*. http://hdl.handle.net/2027/spo.6959004.0001.002.

Monod, J. 1971. *Chance and necessity: An assay on the natural philosophy of modern biology*. New York: Knopf.

Müller, G. 2007. Evo-devo: Extending the evolutionary synthesis. *Nature Reviews Genetics* 8: 943–949.

Nadelson, L.S., and G.M. Sinatra. 2010. Shifting acceptance of evolution: Promising evidence of the influence of the understanding evolution website. *The Researcher* 23: 13–29.

Okasha, S. 2006. *Evolution and the levels of selection*. Oxford/New York: Oxford University Press.

Okasha, S. 2008. Fisher's fundamental theorem of natural selection—A philosophical analysis. *The British Journal for the Philosophy of Science* 59: 319–351.

Pigliucci, M. 2002. *Denying evolution. Creationism, scientism, and the nature of science*. Sunderland: Sinauer.

Pigliucci, M. 2008. Sewall Wright's adaptive landscapes: 1932 vs. 1988. *Biology and Philosophy* 23: 591–603.

Pigliucci, M., and J. Kaplan. 2006. Natural selection and fitness: after the "force" metaphor. In *Making sense of evolution. The conceptual foundations of evolutionary biology*, 13–35. Chicago: University of Chicago Press.

Pigliucci, M., and G. Müller. 2010. *Evolution, the extended synthesis*. Cambridge: MIT Press.

Platt, J.R. 1964. Strong inference. *Science* 146: 347–353.

Prothero, D. 2009. Stephen Jay Gould: Did he bring paleontology to the "High Table"? *Philosophy & Theory in Biology*. http://hdl.handle.net/2027/spo.6959004.0001.001.

Provine, W.B. 2001. *The origins of theoretical population genetics*. Chicago: University of Chicago Press.

Raup, D.M. 1967. Geometric analysis of shell coiling: Coiling in ammonoids. *Journal of Paleontology* 41: 43–65.

Raup, D.M., and S.J. Gould. 1974. Stochastic simulation and evolution of morphology—Towards a nomothetic paleontology. *Systematic Zoology* 23: 305–322.

Recker, D. 2010. How to confuse organisms with mousetraps: Machine metaphors and intelligent design. *Zygon* 45: 647–664.

Ruse, M. 1975. Darwin's debt to philosophy: An examination of the influence of the philosophical ideas of John F.W. Herschel and William Whewell on the development of Charles Darwin's theory of evolution. *Studies in History and Philosophy of Science Part A* 6: 159–181.

Ruse, M. 2000. Darwin and the philosophers: Epistemological factors in the development and reception of the origin of species. In *Biology and epistemology*, ed. R. Creath and J. Maienschein, 3–26. Cambridge: Cambridge University Press.

Saxer, G., M. Doebeli, and M. Travisano. 2010. The repeatability of adaptive radiation during long-term experimental evolution of *Escherichia coli* in a multiple nutrient environment. *PLoS One* 5: e14184.

Sober, E. 2010. Natural selection, causality, and laws: What Fodor and Piattelli-Palmarini got wrong. *Philosophy of Science* 77: 594–607.

Sterelny, K. 2007. *Dawkins vs. Gould: Survival of the fittest*. Cambridge: Icon.

Wade, M.J. 1992. Sewall Wright: Gene interaction and the shifting balance theory. In *Oxford surveys in evolutionary biology*, vol. 8, ed. D.J. Futuyma and J. Antonovics, 35–62. Oxford: Oxford University Press.

Woods, R., D. Schneider, C.L. Winkworth, M.A. Riley, and R.E. Lenski. 2006. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 103: 9107–9112.