The Geological Society of America
Memoir 203
2009

Theory choice in the historical sciences: Geology as a philosophical case study

William L. Vanderburgh†
Department of Philosophy, 1845 N. Fairmount Street, Campus Box 74, Wichita, Kansas 67260-0074, USA

ABSTRACT

Theory choice, the problem of accepting/rejecting scientific theories, is philosophically interesting in part because it involves appeal to nonempirical factors that can only be justified by philosophical considerations. The emphasis in this paper is on the historical as opposed to the experimental sciences—including astronomy, evolutionary biology, and especially historical geology—with examples taken from seventeenth through nineteenth centuries. The fact that evidential reasoning inherently requires a choice of philosophical/methodological principles is demonstrated through reference both to historical cases and to general philosophical considerations. This paper argues that methodological principles play a crucial role in turning empirical data into evidence for/against theories, and it outlines some of the particular evidential and methodological difficulties faced in the historical sciences. Choices of methodological principles depend on nonempirical factors, and because definitive arguments can rarely be found, they are largely a matter of judgment. "Scientific" debates are thus sometimes really disputes over philosophical taste and judgment. Moreover, it is often the case that clear judgments about the incorrectness/correctness of a methodological principle used in a specific context can only be made retrospectively. In part by looking at connections among Isaac Newton, David Hume, and Charles Lyell, and in part by examining Lyell's own arguments, I argue that it was reasonable for Lyell to adopt uniformitarianism as a central methodological principle. Through arguments and historical examples, I also show that there are limits to the acceptability of the uniformitarian position.

Keywords: theory choice, evidence, scientific method, uniformitarianism, Charles Lyell.

INTRODUCTION

When philosophers of science talk about "the problem of theory choice," they have in mind questions about the methods, processes, evidence, reasons, assumptions, and arguments involved in deciding which scientific theories to accept or reject. This paper discusses theory choice using historical geology and other historical sciences such as astronomy and evolutionary biology as illustrations.

The structure of the paper is as follows. Section 1 outlines some of the philosophical considerations arising in theory choice in general. Section 2 elaborates on some details of the role of evidence in theory choice and discusses evidential and methodological problems special to the so-called historical sciences. Section 3

†E-mail: william.vanderburgh@wichita.edu.

Vanderburgh, W.L., 2009, Theory choice in the historical sciences: Geology as a philosophical case study, *in* Rosenberg, G.D., ed., The Revolution in Geology from the Renaissance to the Enlightenment: Geological Society of America Memoir 203, p. 255–264, doi: 10.1130/2009.1203(19). For permission to copy, contact editing@geosociety.org.

focuses on the role of methodological principles in theory choice, with special emphasis on the principle of uniformitarianism in historical geology. Section 4 continues the discussion of uniformitarianism and also considers other principles of method that have been important in the history of the discipline of geology.

ON CHOOSING THEORIES: SCIENTIFIC AND PHILOSOPHICAL CONSIDERATIONS

Investigations into the problem of theory choice ask questions such as these:

1. How do we get from facts to theories?
2. Faced with competing theories, how do we decide between them?
3. What justifies the principles that we do in fact use to make theory choices?

When considering questions such as these, it is important to be aware of the sense in which the questions are asked. If they are asked in a descriptive sense, the answers will have to do with actual practices, either contemporary or historical. If they are asked in a prescriptive (or normative) sense, the answers will tell us what "ought" to be the case, i.e., what the correct thing to do is. As is all too common with human endeavors, what is the case is not necessarily the same as what ought to be the case. It is important to be clear about the two different senses in which theory choice questions can be asked because it is never logically correct to infer an "ought-statement" from an "is-statement" (or vice versa).

Take the question, "How do we decide between competing theories?" There are at least two general categories of answers to consider. One has to do with empirical adequacy, that is, with whether or not the competing theories have equivalent predictive and/or explanatory success. A theory can be described as "empirically adequate" when its predictions agree with the available observations to within the margin of error in the data. Two theories that meet this condition can be described as "empirically equivalent." When a theory fails to be empirically adequate, there are two possible responses: either reject the theory outright, or modify the theoretical background (including background assumptions, descriptions of initial conditions, or parts of the theory itself) so as to make the theory become empirically adequate.

The second category of answers to the question about deciding between competing theories has to do with the methodological features of the competing theories. Most often, explicit discussion of the methodological features of theories as a ground for theory choice occurs when faced with competing theories that are exactly or nearly empirically equivalent. In such cases, theorists will often consider which of the empirically equivalent theories is to be preferred on nonempirical grounds. It is in response to such situations that criteria including simplicity, elegance, explanatory power, consistency with the rest of accepted theory, etc., are most often invoked. Note, however, that while methodological criteria are often applied comparatively, sometimes single theories with no empirically close rivals will be rejected on methodological grounds—for example, when they are thought to be overly complex, inconsistent with the body of accepted theory, inelegant, and so on. Theory choice is, thus, partly a matter of empirical accuracy and partly a matter of methodological adequacy.

To some extent, the relative importance of a particular methodological principle is determined by the reasons for theorizing in the first place. Where one theorizes in order to provide after-the-fact explanations, explanatory power and consistency with other disciplines could be prioritized. Where one theorizes with the aim of making accurate predictions about future observations, simplicity and predictive accuracy could be treated as more important. There are often trade-offs between competing methodological principles, and there is no unique way to "correctly" balance them—this is one of the places where philosophical considerations can have an important bearing on scientific inquiry.

Note that adopting some position or other on methodological/philosophical questions is unavoidable when making theory choices. Most of the time these philosophical commitments are tacit, but without methodological principles, it is impossible to even get started in the basic scientific task of generalizing from finite observed data to unobserved cases. Even a conclusion as simple as "All ravens are black" does not follow immediately from "All observed ravens are black." The leap from the observed to the unobserved must be mediated by a rule for reasoning, a methodological principle, that constrains the possible conclusions that are to be taken to "follow" from the available information. (There is only an imaginary boundary between science and philosophy.)

Although scientists and philosophers often suffer from the illusion that there is a uniquely correct logic of empirical inference, really there are an indefinite number of possible inference rules. Choosing between them—choosing methodological principles—is one of the key functions of the philosophical part of the endeavor to understand the nature of the world in its totality. It is possible to have an inference rule that says, "If all observed cardinals are red, then conclude that all unobserved cardinals are blue." We reject this rule out of hand, but the rejection is rooted in a tacit philosophical position regarding what the correct rule of inference is. Making explicit the philosophical commitments implicit in scientific thinking is one of the useful functions of philosophy of science, in part because it allows the possibility of analyzing whether or not current practices are truly the ideal practices.

Here are some examples of methodological principles of theory choice that should be kept in mind throughout the rest of the discussion. Assuming empirical adequacy, we generally think it is better if a theory is also:

1. simple (as possible; or, simpler than competitors),
2. fruitful,
3. broad,
4. unified,
5. explanatorily powerful,

6. able to make successful novel predictions,
7. consistent with other parts of science,
8. etc.

These and similar methodological principles are the ones normally discussed in the context of the problem of theory choice. Most scientists and philosophers agree that principles such as these are the ones actually used to help guide theory choices. However, is current practice maximally correct? *If* these are the right principles, then employing them correctly will lead to correct theory choices—but why should we think these *are* really the right principles? How can we justify each of those principles/reasons for preferring theories? (Note that I am not seriously doubting that these are the correct principles; I am inquiring into their justification.)

In choosing principles of theory choice, we are ultimately led to ask, "What principles can we use for choosing our principles of theory choice?" This is obviously a difficult question, in part because of the infinite regress of principles it implies (we would need principles for choosing our principles for choosing…our principles of theory choice). It is difficult also because it is a "metascience" question: While the answer determines how to deal with empirical evidence, the question itself cannot be answered by appeal to empirical facts. (It is interesting that methodological principles are, strictly speaking, not based on empirical evidence, yet it is impossible to do evidential reasoning without them.) We might look to examples from the history of science for guidance about which methodological principles have been successful in previous cases, but successful outcomes in the past are no guarantee of the *correctness* of the principles used in the past—the past successes might have been the result of an accidental correlation, a selection effect, or some other sort of evidential illusion. Arguing for or against various possible methodological principles must take place not at the level of empirical claims but at some other (philosophical) level.

It is possible, then, to distinguish three different though related levels of analysis when it comes to the issue of theory choice. (1) Science *uses* methodological principles to make decisions about how to treat evidence and how to evaluate/accept/reject theories. (2) History and sociology of science are interested in the descriptive issue of the rules of theory choice scientists do in fact use, and how those practices have changed over time. (3) Philosophy of science is interested in exploring the prescriptive/normative issue of which methodological principles should be used, and how they ought to be interpreted and applied. Of course, these three levels of analysis need not be isolated, and, in fact, some apparently "scientific" debate is actually normative (that is, philosophical) debate about which methodological principles are appropriate to a given problem. One kind of contribution to that sort of debate is appeal to historical "ideal examples" of scientific theory choice—the narrow sense of a Kuhnian paradigm—as exemplars to guide future practice. Thus, while it is possible to distinguish these three levels of analysis, it would be a mistake to think that they should only, or even can be, pursued independently.

Before moving on, it should be noted that there is disagreement in current scholarship about the possibility of ever successfully addressing the philosophical problems of theory choice. The "logical empiricist" approach to philosophy of science that was the received view for the first half of the twentieth century collapsed under pressure from several directions, including its persistent failure to meet its own standards and research goals. No replacement consensus has yet emerged. This is, thus, a time of ferment in fields related to philosophy of science. An example of the failure of the logical empiricist framework is a persistent inability to find a convincing rational justification of induction; one result of this failure is that we are left with the problem of underdetermination, which says that there will always be an indefinite number of alternative theories available that the empirical evidence will be unable to decide between. (Underdetermination will be mentioned again at the end of this paper; it is one of the central philosophical problems of theory choice.) If no evidential or rational reason for preferring some theories over their rivals exists, the objectivity and rationality of science are called into question. One response to the failure of the logical empiricist program has been pessimism: Social scientists, including some historians but most especially sociologists of knowledge, have given up on the project of coming to a rational understanding of and justification for scientific decision making. They propose, instead, that science is largely irrational and driven almost entirely by sociological and political pressures. This leads to a radical form of relativism. Despite the lack of a consensus about how to replace the logical empiricist framework, many philosophers have remained reluctant to follow this relativist path. However difficult it may be to solve the problems of induction, confirmation, theory choice, and so on, many philosophers prefer to try find new ways to pursue the study of the foundations of science that assume that science is largely objective, rational, and evidentially driven (while recognizing that sociological, political, and other irrational factors nevertheless sometimes play an important role). The current paper falls into this latter category.

EVIDENCE AND THEORY CHOICE IN THE HISTORICAL SCIENCES

In the broadest sense of the term, science is inductive. That is, from limited observed facts, science draws conclusions about unobserved facts. For example, from (present, observed) stratigraphic features, geologists infer the (absent, unobserved) causes of those features. The conclusion of an inductive argument may be a particular or a general claim (e.g., "*this* feature was formed by *this* process" or "*all* features of this type are formed by this type of process"). Inductive conclusions are never certain, only probable (ideally, but not always, highly probable). Any inferential method that draws conclusions that go beyond the data in the premises is inductive in this general sense; I do not intend to limit my claims to enumerative induction. (I should mention here that, following the majority of contemporary philosophers of science and logicians, I include "abduction"—also known as "explanatory inference" or "inference to the best explanation"; see Lipton

[2004]—as a variety of induction. Abductive inferences are inductive in the general sense in that their conclusions make claims that go beyond the information contained in the premises.)

There has long been debate about which particular inductive methods are appropriate to science; the best answer may be that different methods are appropriate to different kinds of needs. Examples of inductive methods include:

1. enumerative induction, which gathers many similar instances and infers the truth of a universal generalization;
2. eliminative induction, which falsifies all but one of the possible alternative hypotheses and concludes that the remaining one must be true;
3. the method of hypotheses, which infers that a hypothesis is true, or probable, because it is consistent with the available evidence;
4. the method of vera causa (real causes), which restricts the method of hypotheses to causes that are known to really exist; and
5. abduction (inference to the best explanation), which considers the possible explanations of some observed facts and concludes that the "best" explanation is the correct hypothesis.

For a discussion of these and some of the many other possible inductive methods, as treated in the context of geology, see Laudan (1982).

Science uses evidence and inductive reasoning to help decide three sorts of questions:

1. What is possible?
2. What is plausible?
3. What is likely?

However, in each kind of case, we can ask what, exactly, is the evidence supposed to be evidence *for*? In geological development as in biological evolution, it is useful to distinguish: (1) the fact of change over time; (2) the path of change over time; and (3) the mechanism of change over time. Michael Ruse uses this tripartite distinction to good effect in his discussion of the evidence for evolution and the debate with biblical creationists (Ruse, 2001, p. 12–32). Applied to historical geology, this distinction yields three questions:

1. Fact Question: Was the Earth different in form in earlier ages?
2. Path Question: What was the sequence of states from earlier ages to the present?
3. Mechanism Question: What were the processes driving the sequence of changes?

Note that these are separate questions, requiring different kinds of evidence for their answers. Also, answers at one level do not necessarily determine answers at other levels. For each kind of question, different kinds of evidence can tell us what is impossible/possible, implausible/plausible and unlikely/likely.

It is clear that different sciences have different kinds of evidence available to them, and that they make different uses of that evidence. A common distinction (somewhat artificial, but useful) is between the "historical" (e.g., cosmology, evolutionary biology) as compared to the "experimental" sciences (e.g., chemistry, population genetics). The difference between historical and experimental sciences is in the kinds of information available from the world. The direct, manipulative studies that are possible in perceptual psychology or high-energy physics are not possible in historical geology or cosmology. The historical sciences have no direct access to their objects of study (past states of a system or set of systems) and hence must infer past states from present traces. Similarly, no direct experiments are possible. Contrast the historical sciences on this score with the physics of falling bodies near Earth, or biological experiments on fruit flies.

That said, "natural experiments" are sometimes possible. Geologists study current volcanic activity in part in order to understand ancient volcanic activity and the structures it created; biologists watch what happens to current populations subject to radical changes of environment in order to learn about long-dead populations.

Finally, indirect manipulative studies are sometimes possible in the historical sciences: that is, direct manipulative studies can be performed on systems that are thought to be analogous to the one of interest. Establishing the degree and the strength of the analogy is crucial to being able to get useful, accurate information from these indirect experiments. For example, we can study stars via the behavior of plasma under artificial conditions on Earth, or we can study natural diamond formation via artificial experiments at high temperatures and pressures. The strength of the evidence obtained by these methods about stars and diamonds depends on the similarities and differences between the objects of interest and the objects actually studied.

Buffon attempted such an indirect experiment when he estimated the age of Earth by comparing its rate of cooling to the rate of cooling of a cannonball (see Gohau, 1990). Buffon reasoned that if Earth started as a molten mass, it must have been impossible for it to have sustained life until it had cooled below a certain temperature. He compared the length of time it took for cannonballs of known mass and volume to cool to room temperature from a white hot state, and on that basis calculated the age of the Earth at around 75,000 years.

This sort of analogical reasoning involves uniformitarian assumptions in two ways:

1. uniformity of Earth's cooling (constant rate over time, etc.), and
2. uniformity of type between Earth and a cannonball.

Unfortunately, both of Buffon's assumptions failed, and hence his estimate of the age of the universe turned out to be wildly inaccurate. (The main problem with Buffon's analogy is that, unbeknownst to him or the rest of science for a long time, radioactivity inside Earth keeps the Earth hotter than it would otherwise be.) Despite its failure (or perhaps because of it), this case illustrates well both the structure of indirect experimental evidence in the historical sciences, and the importance of establishing that the experimental system is appropriately analogous to the historical system about which conclusions are being drawn.

Note that whatever the source of the evidence—whether direct or indirect, experimental or otherwise—the observed data (the quantities and qualities observed) must be turned into evidence. In this process, data are mediated by several factors, including theory, background information, background assumptions (including metaphysical claims about the regularity of nature that are not subject to test), and methodological principles (including a preference for simpler theories, etc.). The process of turning data into evidence can be illustrated by examples from the history of science.

Charles Lyell himself asserts the analogy between astronomy and geology, both with regard to the disadvantageous evidential position and the sameness of the methodological tools employed to overcome that disadvantage.

> It is only by becoming sensible of our natural disadvantages that we shall be roused to exertion, and prompted to seek out opportunities of discovering the operations now in progress, such as do not present themselves readily to view. We are called upon, in our researches into the state of the earth, as in our endeavours to comprehend the mechanism of the heavens, to invent means for overcoming the limited range of our vision. We are perpetually required to bring, as far as possible, within the sphere of observation, things to which the eye, unassisted by art, could never obtain access. (Lyell, 1830, p. 83)

The need to invent new technology—a "geological telescope," if you will, to bring the unseen world into clearer view—is only part of Lyell's meaning here. In astronomy, as exemplified by Newton's great work, *The Mathematical Principles of Natural Philosophy*, new technology only supplied new data. That data had to be manipulated, and Newton's great achievement was in showing how to turn that data (more precise planetary positions, catalogued over time) into evidence about the structure of the solar system and the laws governing gravitational interactions in general. Newton's achievement was partly mathematical, but it was most significantly methodological. In many respects, Newton's methodology became the scientific ideal; with it, he was able to bring into view things such as orbits and mutual attractions that would otherwise never have been visible to the eye—or to the telescope. Lyell is in part proposing a similar recipe for geology: Find those instruments and techniques, including principles of method, which can reliably and plausibly reveal to us the unseen processes that shaped Earth.

To further illustrate some of the methodological issues common to the historical sciences, consider two cases from post-Newtonian astronomy. In 1781, William Herschel (father of John Herschel) discovered a planet beyond Saturn that came to be called Uranus. By 1820, a persistent discrepancy between the Newtonian predictions for and the actual observations of Uranus's position over time was well known. In short, it was impossible to describe a Newtonian orbit that incorporated both the post-Herschel observations of Uranus and the "prediscovery" observations that had been found in various older stellar catalogs (in which the planet's position had been unknowingly recorded as if it were a star). Two possible resolutions of this empirical discrepancy existed. In each case, the "data" are the amount and direction of the discrepancy between the Newtonian predictions and the observed predictions, considered over time. First, via the assumption that Newton's theory of universal gravity is correct, the data become evidence for the existence of a previously unknown mass (John Couch Adams, Urbain LeVerrier). Via the assumption that the known bodies in the solar system are the only gravitationally significant masses, the data become evidence that gravity obeys a non-Newtonian force law at the distance of Uranus (George Biddell Airy, the Astronomer Royal at the time).

A case involving the planet Mercury is an instructive comparison here: The discrepancy in Mercury's motion (also discovered by LeVerrier, in 1843) was of an identical kind, and again both "modification of gravity" and "new matter" solutions were possible. Whereas in the case of the Uranus discrepancy, the correct solution was to predict the existence of Neptune, in the case of the Mercury discrepancy, the correct solution was to develop Einstein's non-Newtonian theory of gravitation. The divergent resolutions of these two exactly analogous empirical discrepancies indicate the difficulty of knowing in advance which of the possible solutions consistent with the data is the correct one. To put the point another way, it is difficult to decide in advance which hypothesis the data are really evidence for.

A similar case can be found in the history of geology. "As early as 1821, Constant Prévost had observed 'a mixture of marine and fluvial shells in the same layers' in the hills of the Paris Basin" (Gohau, 1990, p. 140; quoting Prévost, 1821). These data can be construed as evidence for three different possible explanations:

1. the occurrence of marine invasions at that location,
2. reworking, or
3. the deposits were formed in an estuary.

Although point 1 is the correct interpretation, Prévost applied the principle of uniformitarianism to deny marine invasions and instead advocated the estuary hypothesis. Note, then, that with a different set of background assumptions (perhaps including different methodological principles), the very same data can be taken to be evidence for different hypotheses. This illustration points out one respect in which science is not purely empirical, as popular accounts too often pretend it is.

THE ROLE OF METHODOLOGICAL PRINCIPLES IN GEOLOGY: THE PRINCIPLE OF UNIFORMITARIANISM

It is clear that the principle of uniformitarianism plays a central role in geology today, and in the history of the discipline. What is the foundation or justification of uniformitarianism as a methodological principle? Three sorts of answers can be attempted; none is perfectly satisfactory. First, we can look for direct evidence indicating that in the past the same processes were operating that are observed to operate now. This sort of direct evidence of uniformity is rare at best, not least because of the problems of turning data into evidence mentioned already. Second, uniformity can be entailed by an accepted account of the metaphysics of the universe. The trouble here is that it simply pushes the question of justification back one level: Why *that* account of metaphysics instead of some other? Third, and perhaps most commonly, uniformitarianism can be invoked for reasons of epistemic and/or methodological safety. That is, the uniformitarian principle can be invoked in the absence of reasons to apply a different principle because it is the one most likely to yield results close to the truth (or because no other principle is likely to do better, or it is the principle that is going to stray least often from the truth, or stray by a smaller amount on average, etc.). Of the three available reasons for adopting uniformitarianism, this third one is perhaps the most convincing. Note that this route is still not easy: Some reason for preferring the principle needs to be articulated clearly and needs to be shown to provide adequate justification for accepting the principle.

However, examples of the failure of uniformitarianism in astronomy and geology reveal the limits of a good idea. An example of the failure of uniformitarianism in astronomy is the attempt to solve a discrepancy in the orbit of Mercury that was exactly analogous to the Uranus discrepancy. As in the case of Uranus, matter hypotheses were formulated in the attempt to make Mercury's Newtonian predictions match observation. One after another, these matter hypotheses were shown to fail, either because the matter hypothesized should have been observable but was not, or because the matter hypothesized should have caused additional motions in the orbits of other solar system bodies that were not in fact observed. Ultimately, the Mercury discrepancy was resolved by abandoning the uniformitarian assumption that gravity acts always and everywhere according to the Newtonian description. Instead, Einstein's general theory of relativity successfully accounts for the motion of Mercury without needing to postulate unseen matter, and it shows that Newton's theory of universal gravity is really only accurate when the velocities involved are much less than the speed of light and the gravitational fields involved are weak. At the distance of Mercury, the Sun's gravitational field is strong enough that relativistic effects are important. In other words, the evidence available to Newton (all of it low-velocity, weak-field evidence) does not justify the absolutely universal claim Newton made on that basis: the action of gravity is not "uniform" in the way Newton thought, after all.

In the history of Earth, various examples of the failure of uniformitarianism can also be found. One is the case of the extinction of the dinosaurs; the best explanation available, one that accounts for the pattern of extinction plus the iridium layer at the K-T boundary, is a cometary impact. Of course, that is a very unusual event in present times, one that could not be fairly described as a "normal" event (though we expect that such impacts have happened many times in the past, and that they were much more frequent in the past).

Examples of the failure of respectable methodological principles raise the question: Under what conditions is it appropriate to abandon the assumption of uniformity (or whatever other methodological principles one might care to consider)? To put this sort of question into specific scientific contexts:

1. When is it acceptable to revise a well-confirmed theory of gravity?
2. When is it acceptable to invoke an unusual or unique event in an explanation (e.g., postulating a comet impact to explain a mass extinction)?
3. When and why did it become acceptable to abandon steady-state cosmology?

In more general terms:

1. What constrains nonuniformitarian hypotheses?
2. What makes nonuniformitarian hypotheses reasonable when, most of the time, uniformitarian hypotheses are thought to methodologically superior?

In order to answer these two general questions, several factors need to be considered. The relative empirical success of the best hypotheses constructed under the respective competing methodological principles is crucial. Also, the elegance and simplicity of the explanations derived from the two hypotheses can be important. The only acceptable hypotheses are those consistent with (at least the majority of) other known facts and well-supported theories across the disciplines (chemistry, physics, astronomy, biology, etc.). It would be best if uniformitarian assumptions are abandoned only after several well-constructed hypotheses built on that foundation are shown to be failures. Of course, it is not possible to exhaustively eliminate every possible hypothesis built on the uniformitarian assumption, but a persistent inability to formulate a viable hypothesis within that framework will, generally speaking, make the scientific community more accepting of attempts to adopt a new methodological foundation. (A discussion of Kuhnian "paradigm shifts"—in which old theories and methodological principles are overturned and replaced by new ones—is beyond the scope of this paper.)

Should geology assume uniformitarianism? In general, the answer is yes. The principle is backed by good sense. Often, there is no better principle available, and some principle or other is needed. In addition, theories built on that principle are normally successful.

However, principles such as uniformitarianism can be difficult to interpret and apply to real cases. As Stephan Jay Gould has shown, Lyell's version of uniformitarianism has several parts—uniformity of law, rate, cause, and process—some of which are good and some not. Sometimes, such principles lead us astray (we look in vain for matter solutions to the Mercury discrepancy, say), and, occasionally, despite our initial thoughts, the principles simply do not apply to the cases in which we are interested. This means that there will sometimes arise situations in which it is a mistake to adopt (or persist in) the uniformitarian perspective.

Unfortunately, no general rules exist for determining when it is appropriate to give up strongly held and normally successful methodological principles. Case-by-case analyses, with attention to instructive examples from the history of science, can make us sensitive to the sorts of issues that surround theory choice. The ways in which scientists wrestle with these issues is one of the respects in which science is philosophically interesting.

Buffon's landmark investigation into the structural properties of wood makes an excellent study of the limits of reasoning and the limits of the extrapolation of experimental results to as-yet-unobserved situations. In his study, Buffon showed that structural properties of large pieces of wood (beams, etc.) cannot be predicted merely from the structural properties of smaller pieces of the same type of wood. Buffon recognized the failure of the uniformitarian assumption in this case because he was able to test each of the small-scale and large-scale systems. In the case of the cooling of Earth, on the other hand, Buffon had no choice but to apply the principle of uniformitarianism since he could not actually compare the (unobserved) rate of cooling of Earth to the (observed) rate of cooling of the cannonball.

Is it unreasonable to assume inductive uniformitarianism? David Hume famously showed that induction cannot be rationally justified (Hume, 1999, p. 108–118). Far from rejecting induction on that basis, however, he argued that human beings are psychologically compelled to reason inductively—we are essentially inductive machines, and we cannot change our natures. Moreover, even though he thought it impossible to completely rationally justify induction, he was nevertheless sure that it was appropriate to judge various examples of inductive reasoning as comparatively better or worse (see Vanderburgh, 2005). Taking an evolutionary perspective, one can easily see why animals would end up with the sort of psychological compulsion that Hume supposes we have. If the laws of nature are unchanging over the lifetimes of individuals, projecting past experience into the future is a key to survival (do eat this, do not eat that, avoid this predator, etc.): only individuals who tend to take the past as a guide to the future are likely to survive and reproduce, and hence the inductive tendency will have increasingly high frequency in successive generations. (If the laws of nature are not unchanging over the lifetime individuals, no strategy for making predictions about the future is likely to be successful.)

The upshot is that, even if we cannot fully justify using the principle of inductive uniformity, in the absence of explicit evidence to the contrary, it is both the best available inference rule and we have a strong psychological tendency to employ it. Put it this way: If there is no evidence leading to the conclusion that the uniformitarian assumption is false in a given instance, there is no better (more effective, more reliable) assumption that could be made. This is analogous to the reasoning employed in probability theory and statistics regarding the so-called principle of indifference: In the absence of evidence to the contrary, all possible outcomes are to be treated as indifferently (that is, equally) probable. (Unless there is some reason to think the die is unfair, treat each number as equally likely to come up on a given roll.)

Does the assumption of uniformitarianism make it impossible to admit nonuniform causes into science? Not at all. Cosmologists recognize that we can project the current laws of nature backward in time only so far: In the conditions of temperature and pressure thought to be present in the earliest moments of the big bang, our current understanding of particle physics and gravitation breaks down. Similarly, geologists can reach the conclusion that in the earliest phases of Earth's development, the intensity of geological processes, if not the kinds of geological processes acting, could have been different. Similarly, nonuniform events such as comet impacts can be accepted as occurring in the past. It is true, however, that claims about unusual events require extraordinary evidence. The methodological assumption of inductive uniformity can reasonably be put aside only when the evidential situation demands it.

MORE METHODOLOGICAL PRINCIPLES IN GEOLOGY

One of the problems with the geological record is that its archives are "defective": they are incomplete, incompletely known, faulty, biased, etc. This is a feature of the available evidence that is paralleled in astronomy, evolutionary biology, and other historical sciences. Lyell in many places uses the word "monuments" to refer to geological features that are interpreted as records of ancient geological processes. The analogy to human historical monuments is deliberate, and interesting. Human monuments are not identical with the events they memorialize, but they "stand for" and record those events. When someone comes across a historical marker (especially when archaeologists encounter them), a great deal of interpretation is normally required in order to make the monument "speak." The information gleaned from such monuments is sometimes biased, faulty, incomplete, or otherwise problematic. The reliability of these interpretations and extrapolations is improved when a greater amount of historical context is known, and when multiple, independent markers provide the same information. Clearly, all of these ideas apply to inferences made from geological monuments too. One of Lyell's main methodological thoughts about the interpretation of geological monuments is that "a considerable part of the ancient memorials of nature were written in a living language" (Lyell, 1830, p. 73); this is to say that contemporary evidence about geological and biological systems is relevant to constraining interpretations of ancient geological events and helps to make those interpretations reliable. (As a referee for this paper pointed out, one difference between typical historical monuments and the geological record, of course, is that historical monuments are not usually the direct result of the events they memorialize, whereas the geological record is directly and genetically related to the specific processes that caused the record. See Baker [1999] for more on the "semiotic" interpretation of geological reasoning. The methodological point I want to make stands despite the fact that Lyell's analogy is imperfect in this way.)

Other methodological principles have been important in the history of the discipline of geology. Nicolas Steno (1638–1686) is famous for his laws of stratigraphy; they are principles that govern and constrain the "allowable" inferences from observed geological data to theoretical generalizations about the formation of geological structures. The principle of superposition demands that strata be interpreted in such a way that younger strata are always originally formed above older strata. The principle of initial horizontality, similarly, governs theorizing about the formation of structures that are not now horizontal. The principle of strata continuity allows one to assume that strata continue laterally far from where they presently end. The principle of crosscutting relationships says that something that crosscuts a layer is younger than that layer (Steno, 1968 [1669]). Note that these principles, which now seem so natural to geologists, needed to be accepted before inferences from geological data could even begin.

Some of the methodological principles that have been attributed to James Hutton (1726–1797) include: geological time is limitless; the present is the key to the past; the internal heat of Earth drives geological processes; the lifecycle of continents involves erosion, deposition, and elevation; geological structures are built from a succession of small events over long periods, not by sudden, brief events/catastrophes; geology should be integrated with the other sciences.

"Some areas of substantive agreement [between mythical foes Hutton, a Plutonist, and Werner, a Neptunist] were much wider than one would imagine, but differences of scientific method, of the proper scope and aim of geological science and of the relationship between fact and theory, made it difficult for the antagonists (or indeed the mediators) to achieve anything like consensus" (Greene, 1982, p. 29). This is to say, then, that the debate between the Huttonians and the Neptunists was at least in large part a debate over which methodological principles were the best ones to apply in geological theorizing. One might even say that the debate was philosophical rather than geological.

Charles Lyell (1797–1875) critiqued Hutton's reliance on internal heat and his catastrophic theory of upheaval. This could only have been a philosophical difference at the time, since Earth's internal heat was not then well understood, and since the actual history of Earth is even now a matter of inference and speculation rather than something directly observable (that is, Lyell could not have claimed that we have direct empirical evidence that catastrophes have never occurred).

Philosophical debate is not the only way in which methodological disputes get resolved. Sociological and historical factors often play important roles as well. For example, Lyell's methodology came to dominance in part because Charles Darwin read Lyell's *Principles* on the Beagle—the success of Darwin's theory of evolution conferred after-the-fact credibility on Lyell later. Similarly, the popularity of Abraham Gottlob Werner (1749–1817) as a teacher was in large measure responsible for the ascendancy of his Neptunist views, despite evidence contradicting some his most significant claims about geological facts and their causes. Lyell bemoaned this "retrograde movement" in the discipline of geology, and used the opportunity of discussing it to advance his own uniformitarian methodology against that of Werner's Neptunian disciples.

> His theory was opposed, in a two-fold sense, to the doctrine of uniformity in the course of nature; for not only did he introduce, without scruple, many imaginary causes supposed to have once effected great revolutions in the earth, and then to have become extinct, but new ones also were feigned to have come into play in modern times; and, above all, that most violent instrument of change, the agency of subterranean fire. (Lyell, 1830, p. 58)

One of the key things to notice about this debate, as characterized by Lyell, is that it is fundamentally a methodological debate. It is not that the evidence itself proves or disproves any of Werner's hypotheses—to use one of Lyell's examples (Lyell, 1830, p. 60), Werner's hypothesis that obsidian is an aqueous precipitate. Rather, the best interpretation of the evidence—the best methodological stance toward the evidence—gives reasons for or against the hypotheses. In particular, Werner's school adopts a methodological principle for interpreting geological data (namely, that all ancient rocks have a sedimentary origin) that Lyell rejects. Lyell rejects it because it leads to conclusions that violate the principles of good analogical reasoning. When Lyell calls Werner's theory of the origin of trap rocks "one of the most unphilosophical [theories] ever advanced in any science" (Lyell, 1830, p. 59), he means that accepting that theory involves bad judgment, in particular bad judgment about which methodological principles to adopt.

One of Lyell's arguments in favor of adopting the uniformitarian framework is this:

> We have seen that, during the progress of geology, there have been great fluctuations of opinion respecting the nature of the causes to which all former changes of the earth's surface are referrible. The first observers conceived that the monuments which the geologist endeavours to decipher, relate to a period when the physical constitution of the earth differed entirely from the present, and that, even after the creation of living beings, there have been causes in action distinct in kind or degree from those now forming part of the economy of nature. These views have been gradually modified, and some of them entirely abandoned in proportion as observations have been multiplied, and the signs of former mutations more skilfully [sic] interpreted. Many appearances, which for a long time were regarded as indicating mysterious and extraordinary agency, are finally recognized as the necessary result of the laws now governing the material world; and the discovery of this unlooked for conformity has induced some geologists to infer that there has never been any interruption to the same uniform order of physical events. The same assemblage of general causes, they conceive, may have been sufficient to produce, by their various combinations, the endless diversity of effects, of which the shell of the earth has preserved the memorials, and, consistently with these principles, the recurrence of analogous changes is expected by them in time to come. (Lyell, 1830, p. 75)

Lyell's initial claim here about the history of the discipline of geology is more or less uncontroversial: many early geologists did invoke causes that are not analogous to any causes currently acting (for example, global deluges and conflagrations). A skeptical reader might, however, doubt his claim that the natural progression of the discipline led to the recognition that the geological evidence was more consistent with uniformity than catastrophe. That claim has the flavor of a rhetorical maneuver to support his own uniformitarian view by stating it as if geologists in general arrived at that view by examining the geological evidence itself.

There are other, better, arguments for uniformitarianism, however. For example, Lyell asks us to consider how the monuments in Egypt would have been interpreted had we held the belief that Egypt had never been occupied by humans until modern times, in the same way that some people held that the Earth was never populated by living beings until the continents were in their present positions. He points out that as new discoveries were made in Egypt, the myths needed to explain them on this hypothesis would have become more and more fanciful ("visionary" as in seeing a vision, an illusion):

> Each new invention would violate a greater number of known analogies; for if a theory be required to embrace some false principle, it becomes more visionary in proportion as facts are multiplied, as would be the case if geometers were now required to form an astronomical system on the assumption of the immobility of the earth. (Lyell, 1830, p. 77)

In short, the greater the number of disparate facts known, the wilder are the ad hoc maneuvers needed to maintain a theory founded on a false assumption. The good sense of the scientist will in such circumstances judge that the best route is to reject the assumption that makes the ad hoc modifications necessary and replace that assumption with another that makes the system of beliefs more harmonious. Lyell here argues, then, that the best interpretation of the evidence comes when we reject the assumption that life has only recently arrived on Earth.

Note that the date of the origin of life is not something about which Lyell has direct evidence. The whole business depends on making "reasonable" judgments about the methodological principles to adopt in a given evidential context. As Pierre Duhem, V.W.O. Quine, and other philosophers have shown, there is no unique way to do this; it will normally be possible to come up with several different sets of initial assumptions, where each set is equally empirically good at explaining the known data. This is called the "underdetermination of theory by evidence" (see Laudan, 1990, and Laudan and Leplin, 1991). Lyell proposes that his own methodological principles, including uniformitarianism, make the most reasonable story out of the data. What recommends the Lyellian principles, he believes, is that they allow us to form a less fanciful, more consistent picture. Although there may be disagreement about the methodological principles that lead to the most coherent and least ad hoc story, even today, consistency (internal to the theory, and across theories) remains the fundamental standard for assessing theories in the historical sciences. Pierre Duhem (1982 [1914]) said that such assessments must be left to the "good sense" of the experienced scientist. Although many philosophers have been dissatisfied with this conclusion and have hoped to come up with objective, rational principles for deciding which scientific judgments are actually well founded and which are not, such principles have remained elusive. So far, the best we can say is that in careful case-by-case analyses of instances in the history of science, there often does seem to be retrospective scientific and philosophical consensus about which judgments are good and which are not. Whether or not the lessons of history will be a useful guide in current scientific theorizing remains to be seen.

CONCLUSION

Methodological principles such as parsimony, explanatory power, unifying power, etc., play a crucial but rarely acknowledged role in all scientific theory choice problems. This is so in part because methodological principles are used as tools for getting around gaps and errors in the evidence, and in part simply because every problem of inductive reasoning is (by nature) unavoidably underdetermined by the evidence. An important consideration, then, is how to choose and evaluate methodological principles. It has been argued here that reason and argument can provide grounds for choosing and evaluating methodological principles, but that, since it is rare to find definitive reasons in this area, much remains a matter of philosophical taste and judgment. One of the useful functions of philosophy of science is revealing and providing a context for discussing these issues in science.

Historical examples from geology and other so-called historical sciences may be used to illustrate the ways in which methodological principles come into scientific theory choice and the difficulties surrounding this. Given the focus of this volume on the rise of scientific geology, emphasis was given to Charles Lyell's use of methodological principles, particularly his uniformitarianism. Examples and arguments show that while the uniformitarian hypothesis has good grounds, it is not always the correct principle to apply. Again, this means that philosophical taste and judgment play a crucial role. This, in turn, helps to explain some of the cases of radical theory change observed in the history of science (some paradigm shifts are the result of shifts of philosophical taste), and it also helps to explain why some "scientific" debates (e.g., Plutonists v. Neptunists) are so acrimonious: The antagonists disagree about deeply felt but difficult to defend philosophical principles. Consensus on these topics is normally only found retrospectively.

It is important to point out that Lyell makes a point that is very similar to the point being made today by scientists and philosophers who disagree with intelligent design theory. Features of the world having origins that we cannot at present understand should not automatically be attributed to the agency of a supernatural power. "God did it," is not an explanation at all, let alone a scientific explanation. Rather, we should notice that we have had good success in finding explanations within the framework of known science for facts that were previously not understood. It is, thus, methodologically better to assume that the next time we encounter some fact that we cannot explain, its explanation will eventually be found within known science, without the need to appeal to unknown singular causes or changing laws. This might turn out to be mistaken, but it is the best place to start.

During the Scientific Revolution and Enlightenment, discussion of methodological principles and correct procedures for acquiring knowledge was at the forefront. We see this as much in philosophers such as Descartes and Hume as in scientists such as Newton, Lyell, and others. It is true of science as a whole as well as for particular sciences. These writers recognized that acquiring knowledge through empirical inquiry requires making decisions about methodology. Through a long period of ferment up through the Enlightenment, consensus was reached on the core of what is now called the scientific method. We now accept, for example, that theories ought to be predictively successful, explanatorily powerful, parsimonious, testable, falsifiable, and so on. At the frontiers of new science, there is less consensus about method, and the details of how to conduct science are still debated. This includes ideas about what kinds of things count as evidence, and about the kind of weight different evidence has in theory choices. These issues will eventually be resolved by making philosophical/methodological choices, in short, judgment. An awareness of the philosophical issues surrounding theory choice, and of historical examples of theory choice, could well contribute to future scientific progress. It has certainly done so in the past.

REFERENCES CITED

Baker, V.R., 1999, Geosemiosis: Geological Society of America Bulletin, v. 111, p. 633–645, doi: 10.1130/0016-7606(1999)111<0633:G>2.3.CO;2.

Duhem, P., 1982 [1914], The Aim and Structure of Physical Theory (trans. Philip P. Wiener from the second French edition): Princeton, New Jersey, Princeton University Press, 344 p.

Gohau, G., 1990, A History of Geology: New Brunswick, New Jersey, Rutgers University Press, 276 p.

Greene, M.T., 1982, Geology in the Nineteenth Century: Changing Views of a Changing World: Ithaca, New York, Cornell University Press, 328 p.

Hume, D., 1999 [1748], An Enquiry Concerning Human Understanding (Beauchamp, T.L., ed.): Oxford, Oxford University Press, 456 p.

Laudan, L., 1990, Demystifying underdetermination, *in* Wade Savage, C., ed., Scientific Theories: Minneapolis, University of Minnesota Press, Minnesota Studies in the Philosophy of Science, v. 14, 433 p.

Laudan, L., and Leplin, J., 1991, Empirical equivalence and underdetermination: The Journal of Philosophy, v. 88, p. 449–472, doi: 10.2307/2026601.

Laudan, R., 1982, The role of methodology in Lyell's science: Studies in History and Philosophy of Science, v. 13, p. 215–249, doi: 10.1016/0039-3681(82)90009-7.

Lipton, P., 2004, Inference to the Best Explanation (2nd edition): London, Routledge, 240 p.

Lyell, C., 1830, Principles of Geology, Volumes 1–3: London, John Murray, 460 p. Available at: http://www.esp.org/books/lyell/principles/facsimile/title3.html (last accessed 7 May 2008). (Electronic Scholarly Publishing, prepared by R. Robbins.)

Prevost, Constant, 1821, Sur un nouvel exemple de la reunion de coquilles marines et de coquilles fluviatiles dans les memes couches: Journal de Physique, v. 92, p. 418–428.

Ruse, M., 2001, Can a Darwinian be a Christian?: Cambridge, Cambridge University Press, 254 p.

Steno, N., 1968 [1669], The Prodromus of Nicolaus Steno's Dissertation Concerning a Solid Body Enclosed by Process of Nature within a Solid (trans. J.G. Winter, with notes): New York, Hafner Publishing Company, 203 p.

Vanderburgh, W.L., 2005, Of Miracles and Evidential Probability: Hume's 'Abject Failure' Vindicated: Human Studies, v. 31, p. 37–61.

Manuscript Accepted by the Society 1 October 2008