INFERENCE TO THE BEST EXPLANATION

by Philip Clayton

Abstract. The common role of research programs in science and religion is now widely accepted. The next step in the methodology debate is to specify more concretely the shared standards for adequate explanations. The article presents a detailed account of the method of inference to the best explanation and gives examples of how the method can structure the philosophical and theological interaction with science. The resulting approach dispenses with deductive and inductive proofs of religious propositions and limits itself to initially plausible hypotheses that are to be assessed according to their explanatory power. Only when a domain of data and a particular explanatory task have been specified can any serious claim be made that religious theories are equal or superior to their naturalistic alternatives.

Keywords: inference to the best explanation; Imre Lakatos; Peter Lipton; philosophy of science; rationality; science and religion; scientific method.

A number of philosophers, while admitting of no compelling deductive argument for God's existence, consider religious belief to be justified on the basis of its power to explain some phenomenon or phenomena. Theism, on this account, is an explanatory theory, or hypothesis, whose acceptability is measured by its explanatory power.

—Robert Prevost, *Probability and Theistic Explanation*

Let us begin with an example. We know that we are the result of an evolutionary process, the general principles of which apply to all living organisms. Genetic variation and selection by environmental forces provide the best explanation for the biological data available, but we also observe significant *discontinuities* between living organisms; moreover, there seem to be properties that emerge only at the level of higher organisms, and there are some distinct properties that one finds only in human beings. Is a *random* or unguided process of variation

Philip Clayton is Associate Professor of Philosophy at California State University, Rohnert Park, CA 94928.

[*Zygon*, vol. 32, no. 3 (September 1997).]
 ISSN 0591-2385

starting from an initial group of cells the best way to account for the variety of species that we encounter on this planet? Or would (for instance) a *nonrandom* process, or separate starting points for major kinds of life forms, better explain the biological data? How shall we decide? And how can we defend our decision procedure in a way that does not beg the question on behalf of either naturalism or theism?

A recent development in epistemology and the philosophy of science[1] has important implications for answering these questions. The goal of this essay is to show in what sense inference to the best explanation (IBE) represents the correct description of the task in *meta*-physical research as well as in physical-scientific work. Using this method, I shall advance specific proposals for how the dialogue between religion and science is to be understood, how it should be pursued, and what results it may attain.

I

The first presupposition for such a discussion is the failure of deductive proofs for the existence of God. Both for general epistemological reasons and because of specific difficulties with the concept of God, it has turned out not to be possible to deduce one's way from undeniable premises to God's existence (see Clayton 1996). However, an argument that fails to prove—that is, to compel the reader to a specific conclusion—*can still constitute a better explanation of some domain than any of its competitors*. For instance, it is possible that our universe could have come to exist without a transcendent, self-caused cause; hence, it is not logically necessary, as the cosmological proof alleges, that there be a creator who is responsible for creating all matter in the universe. By reformulating the argument, however, we can inquire whether the hypothesis of theism provides a good explanation for the matter and types of organization that we do find. In other words: given an adequate theory of explanation, *God* can now be taken as a theoretical term in an explanatory hypothesis and not as a derivation from first principles.

The debate over the anthropic principle(s) in recent frontier discussions between science and religion provides another good example. The problem is that the strong anthropic principle is often taken as a quasi-deductive argument for the existence of God:

1. Precisely the right conditions existed to allow for the emergence of human life;
2. The probability of this precise combination of conditions' occurring on its own, without intentional intervention from some source, is virtually nil; hence,

3. An extremely powerful conscious being must exist who willed that human beings exist and who constructed a natural order that would (probably) lead to their emergence; that is,
4. God exists.

Powerful criticisms of this argument have been raised, some of them in the pages of this journal. Should we then dismiss anthropic considerations as worthless? Here again, I suggest, the actual evidential weight of the data is best understood as an inference to the best explanation. The question is, Given the existence of exactly these values, together with the high initial improbability of exactly this sort of organization, what best explains the data? In order to come finally to a comparative weighting of the most serious contenders, one first works to specify the exact explanatory question at hand and then draws up a list of viable hypotheses. Of course, because there are a number of live options, the result will not be a 100-percent weighting for theism—or for naturalism!—but rather a weighting across the spectrum, according to which carefully nuanced comparative assessments can be made.[2] Under this interpretation, philosophical arguments do not prove the existence of God but, rather, show "the explanatory power of the concept of God by formulating fundamental questions to which the existence of God could constitute an intellectually satisfying answer" (Hick and Goulder 1983, 33).

Or consider another example: the theory of common ancestry (TCA). A rigorously *naturalistic* theory of the emergence of life on earth requires TCA if naturalistic evolution is to provide, even in principle, a full explanation of the life forms that we find. (It also requires, among other things, a credible account of the transition from inorganic matter to primitive cells and an adequate selection-based explanation of the development of complex organs.) By contrast, theories of *theistic* evolution are compatible in principle either with TCA or with multiple creative interventions on the part of God (although I suppose that the more natural reading of Genesis would be with a series of separate creative acts by God). The two positions are not incommensurable, and a comparative evaluation is possible via the mechanisms of inference to the best explanation.

Clearly, in disputes such as TCA no deductive argument can be made from the scientific data to either naturalism or theism. What happens when we try to understand such debates as *inductive* arguments? According to the best contemporary theory, one assesses the respective inductive probabilities of two or more explanations using the Bayesian calculus.[3] For a given hypothesis h, body of evidence e, and agreed-upon background evidence k (where k must be independent of e), the probability of the hypothesis is stated as $P(h \mid ek)$. This probability is calculated using Bayes's theorem: $P(h \mid ek) = P(h \mid k) \; [P(e \mid hk \mid P(e \mid k)]$. Separate

probability assessments are developed for the two competing theories, and whichever has the higher value (the value approaching 1.0 or certainty) is considered to be rationally indicated.

But I suggest that Bayesian probability as a tool for assessing theistic claims is beset with serious difficulties. Perhaps the most serious one is that theories such as TCA have different initial probabilities, depending on whether the background assumptions are theistic or naturalistic. Generally, for any given naturalist (N) and theist (T), the initial subjective probability of TCA should be higher *relative to* N's background assumptions than to T's. Given N's background assumption that divine interventions in history are unlikely, and the fact that naturalistic evolution requires TCA, N correctly ascribes a higher initial probability value to TCA. T is equally rational in using the assumption of the likelihood of divine intervention to ascribe TCA a *low* initial probability. The disagreement between N and T cannot be resolved by appeal to the evidence *e*, because Bayesian calculus requires that *e* be independent of the background assumptions. Given this stalemate, we might decide to turn away from evolution for a moment and try first to come to agreement on the initial probability of divine intervention in history. Then, however, I suggest that we would find ourselves *either* with diverging interpretations of the evidence itself, and hence without the possibility of resolution, or with diverging background assumptions, which would themselves have to become the object of further inquiries and probability assessments, and so on ad infinitum. The reason for the failure of the Bayesian method in such questions is that fundamental background issues (*k*) are *essentially contested* (see Gallie 1995–96); and where *k* cannot be determined, the probability calculations cannot be made.

Clearly, these are complicated questions, and the foregoing does not offer a decisive refutation of deductivist or inductivist natural theology. The goal here, instead, has been merely to encourage the suspicion that probability assessments regarding decisive disputes between naturalism and theism turn out to rely on background questions about which naturalists and theists do not agree. Hence, the probability calculus cannot itself be used to resolve those disagreements, at least not without the threat of vicious circularity. But if deductive theistic arguments fail and Bayesian probabilities are not helpful, it is doubly urgent that we develop a mechanism for assessing theistic proposals. The theory of inference to the best explanation provides, I propose, the best framework for this task.

II

According to Inference to the Best Explanation, we infer what would, if true, be the best explanation of our evidence. . . . We have to produce a pool of potential explanations, from which we infer the best one. . . . One of the main attractions of

> the model [of Inference to the Best Explanation] is that it accounts in a natural and unified way both for the inferences to unobservable entities and processes that characterize much scientific research and for many of the mundane inferences about middle-sized dry goods that we make every day. (Lipton 1991, 59, 22, 188)[4]

We have already seen the advantages of viewing the key issues on the borderline between religion and science primarily in terms of explanation. Negatively, we have found that the alternative views—the deduction of religious truths and strictly inductive arguments for the probability of God, not to mention attempts at decisive falsification of theological hypotheses (Lipton 1991, 94)—are not successful. Positively, the explanatory quest, with its attempt to move back through the chain of why's, nicely captures at least one major function of (and motivation for) talk of God (Clayton 1989, e.g., chaps. 1 and 5). There are, of course, a variety of theories of scientific explanation, including the reason model, the familiarity model, and the deductive-nomological model. Although we cannot canvass them all here, I suggest that the difficulties the other models face may be serious enough to warrant a careful look at this competing model (cf. Clayton 1989, 26–31; Lipton 1991, chap. 2).

Unlike the other theories, inference to the best explanation correctly grasps the fundamentally *contrastive* nature of explanation. In any attempt to explain, "what gets explained is not simply 'Why this?', but 'Why this *rather than* that?'" (Lipton 1991, 35).[5] We could thus speak of the components of explanation as involving a *fact to be explained* (an explanandum) and a *foil* (something that might have been the case but is not). A successful explanation always shows why a specific fact exists and the foil does not. Different foils may help to pick out and explain different parts of the same causal structure of the world (Lipton 1991, 76). It follows immediately from this crucial criterion that those discussions are explanatorily vacuous in which the foils are not clearly specified. So in order to initiate a rational debate in cosmology, one cannot merely ask, for instance, "Why are there human beings?" Rather, the question will be Why is this amount of order to be found in the universe rather than less order (or more)? or Why are intelligent animals present on this planet rather than only lower life forms (or none at all)?

The foil criterion in inference to the best explanation actually turns the testing of explanations into a two-step process (Lipton 1991, 60 ff.). One might say that there are two "filters:" an initial decision regarding plausibility and a final decision to select the best of the competing explanations. Before the actual competition for the best explanation, one must *already* have generated a set of plausible candidates, a pool of "live options." We do not posit the influence of phlogiston in searching for new explanations in chemistry, or ether in physics, nor should we start

with the possibility of mass hypnotism by alien beings when we try to explain ups and downs in the stock market. Plausibility judgments are not completely subjective or arbitrary, as the foregoing examples show; there has been real theoretical progress in science in the sense that certain explanatory options have become less plausible with time. The same can perhaps be said of theology, although in this case it is probably more accurate to speak of an increased awareness of the *difficulties* that particular views face, such as the Scholastic view of God as in no way dependent on the world or the premodern view of God as continually breaking the laws of the natural order that he created.

Choosing a foil, however, is also not a purely objective process, as Bas van Fraassen has argued. Different factors are salient for a person because of "his orientation, his interests, and various other peculiarities in the way he approaches or comes to know the problem—contextual factors" (van Fraassen 1977; 1980, 125). Because explanatory interests vary, initial assessments of plausibility cannot be viewed as those that a neutral observer might determine, even in principle, for all explainers. This was the reason for moving beyond the probability calculus in the first place: naturalists and theists inevitably (and with good reason) weigh initial probabilities differently, at least for issues of contention between them, such as miracles, the likelihood of living beings' emerging from inorganic matter, or the need for an infinite or transcendent cause of the Big Bang. The best way to get a handle on this difference in interest, according to inference to the best explanation, is to specify the competing interests as distinct foils so that we ask, "Why *P*?" in contrast to specific other alternatives *Q* of a given set *X* of options.[6]

Once a particular explanatory contrast is specified and initial plausibilities are determined, we can then make progress using the method of "causal triangulation." Like J. S. Mill's "Method of Difference," this method expresses the requirement that "a [specifiable] cause must lie among the antecedent differences between a case in which the effect occurs and an otherwise similar case where it does not" (Lipton 1991, 43). For example, if we find a significant difference in the performance of students at two schools, we may begin by trying to eliminate dependent variables, such as bad lighting or inferior textbooks. As the learning environments (and hence the causal descriptions) of the schools become more and more similar, the number of possible explanations for performance differences is reduced. This process of triangulation finally makes it possible to hypothesize one or more specific causes of the difference (teaching methods, say, or the children's nutrition), which we are then justified in advancing as the best explanation of the difference.

The focus on comparative histories or causal stories offers an important corrective for meta-physical discussions. Scientists almost always tell

causal stories: Why did the maple trees die? Well, because the acid rain caused such and such a physiological state, and trees in this state can no longer live. The sense of dismay that my scientific friends often express regarding religious explanations is that they cannot identify precisely what causal claims are being made. Religious "explanations" of scientific results are like tennis without a net, they complain: any phenomenon whatsoever seems equally amenable to detailed religious appropriation and explanation.

The fact-foil requirement for an explanation can help to minimize such differences between scientific and religious explanatory accounts. It demands that religious explainers show that theirs is the best explanation for why one particular state of affairs is to be found rather than some specific alternative. Speculations about the nature of God are controlled (somewhat) by the requirement that one specify the particular characteristics of the physical world, *rather than some other set of characteristics*, which one's theology is designed to explain. Religious theories of value should likewise be required to list the particular values (or characteristics of humans as valuing animals) that they rely on as data, along with the alternatives that might have been but are not. This exercise can also have negative results. Theologians have argued, for example, that those physical constants necessary for sustaining human life that we have measured have these values *rather than others* because the universe was designed by God so that intelligent beings would evolve (the teleological argument). The foil draws attention to a weakness in this explanation, however: if the constants had indeed had other values, human beings would not be here to observe the universe at all. Careful consideration of the foil makes it less clear that a cosmic designer is really needed to explain our discovering these constants; they are already presupposed (and in that sense, explained) by our existence in the first place.

Is the compatibility of scientific and religious explanations ensured by the foil technique? Not yet. Recall that the theory of inference to the best explanation "suggests . . . that the central requirement for a sensible contrastive question is that the fact and the foil have a largely similar history, against which the differences stand out. When the histories are too disparate, we do not know where to begin to answer the question" (Lipton 1991, 46). Prima facie, it appears that causal histories of the world in strictly naturalist terms and in terms of God's creative agency are as divergent as one might wish! Only if we can bring the two accounts into contact can we reduce the disparateness and begin to speak of compatible projects. (Of course, this presupposes that we are able to give a coherent account of what we mean by *divine agency* in the first place.) One (and perhaps the only) way to establish compatibility is to develop a more general theoretical framework, one

that is neither purely scientific nor purely theological and thus can bring together the two different types of causality, allowing us to tell a single causal story. It seems to me that serious proposals in our field —Pierre Teilhard de Chardin's theory of evolution, or Frank Tipler's notion of the Omega point as "eschaton," or Philip Hefner's idea of created cocreators—should and can be judged by the extent to which they are able to establish a real causal overlap with the scientific theories to which they make reference. This is not to imply that there will then be *no* differences from the scientific accounts; in most cases a difference of explanatory interests will remain. Still, if our "dialogue" with science is to be genuine, the novel inferences made in religious explanations of scientific data should depend on *relevant* differences (Lipton 1991, 80); and it is incumbent on us to *show* the relevance of our explanatory hypotheses to the sciences in question.

Perhaps it will seem to religious readers that I am overemphasizing the burden of proof faced by religious explanations. This is not the case, however. Careful, nonpositivistic descriptions of science, such as Lipton's work on inference to the best explanation, although they do not rule out theistic hypotheses, do reveal what it is that gives science its incredible explanatory success. In scientific research one finds careful controls on the introduction of unobservables (Lipton 1991, 89); explanations can appeal to "causal histories" that are accessible to observation and control; careful experiments can be designed to eliminate alternative explanations (Lipton 1991, chap. 5); and risky predictions guard against possible "fudging" on the part of theorists working to accommodate their data (Lipton 1991, chap. 8). If we are to maintain the rationality of religious theorizing, we must maximize the opportunities for "the feedback between the processes of hypothesis formation and data acquisition that characterizes actual inquiry" (Lipton 1991, 89). Gone, of course, is the old demand for strict falsifiability, which was as much a chimera in science as in theology. Instead, according to inference to the best explanation, disconfirmation comes about through evidence that (for example) simultaneously strengthens a theistic proposal and weakens the competing naturalistic alternatives. Nonetheless, there is a scientific *attitude* that is sorely missing in many of our more theological discussions. The scientific preference for prediction stems directly from the sense that "we should test the scientist, and that she should test herself" (Lipton 1991, 154). Only when this attitude is noticeable in meta-physical explanations, and when it begins to bear appropriate fruit, can we begin to claim epistemic equity with our scientific colleagues and to enjoy the laurels thereof.

Let us suppose, then, that inference to the best explanation provides not only the most adequate model of explanation in science but also, as I think, a viable methodology for philosophers and theologians. It now

becomes urgent to specify what constitutes the "best" explanation among a set of plausible candidates. Before we can begin to list specific criteria to use in a given debate, we must specify in more general terms what we mean by *best*. There are two major options: "the explanation best supported by the evidence" and "the explanation that would provide the most understanding," or, in Lipton's terms, the *likeliest* and the *loveliest* explanations (Lipton 1991, 59). There is an obvious problem with the former, however: if we are searching for the best explanation, the one that has the strongest claim on truth, it is unhelpful to be told that we should seek the likeliest candidate—that is, the one most likely to be true! Hence, the thesis: We are rationally justified in selecting the loveliest explanation, the one "which would, if true, provide the most understanding" of the domain in question (Lipton 1991, 186).

What is a loveliest explanation? It is one that instantiates the *explanatory virtues*. (Note the shift away from direct talk of *truth criteria*: although the goal is still the same, philosophers of science more often begin now with talk of the epistemic values underlying the explanatory quest. At best, it seems, we can derive a few general desiderata, which will never be sufficient for deciding between theories; they will have to be supplemented by other criteria that are relative to specific disciplines, situations, and sets of theoretical interests.) Among the explanatory virtues are aesthetic considerations such as theoretical elegance, simplicity, and coherence (Lipton 1991, 68). Precision is another—we generally choose an exact explanation over a vague one—and a third is "mechanism" (Lipton 1991, 118). Scientists prefer explanations that specify the causal mechanisms that lead to the observed effects. Finally, we value a "unified explanatory scheme" (Lipton 1991, 119), one that can fit a given phenomenon into the broadest possible theoretical structure.

This last criterion tells both for and against religious or metaphysical explanations. By their very nature such explanations are generally broader than the scientific discipline with which they make contact; hence, the odds are good that they will be able to unify partial explanations arising in particular scientific domains. As Lipton correctly points out: "When a theory provides a unified explanation of many and diverse observational judgments, and there is no remotely plausible alternative explanation, we may have more confidence in the theory than we had in the conjunction of the evidence from which it was inferred" (Lipton 1991, 182 f.).

In both theological cosmology and theological evolution the typical move is to encompass scientific results within an overarching theory of God's purposes and actions in the world. Yet along with this breadth comes a disadvantage: the theologian must introduce *types* of causes— divine agency, God's purposes, the lure of The Good—which are disanalogous to the ordinary causal forces in the world. Theological causes are often

characterized as *transcendent*, unique and without parallel to any other domain of our experience. What such accounts gain in explanatory comprehensiveness they run the risk of losing through the uniqueness claim, which makes them difficult to conceptualize, to identify, and to check.

III

What does the theory of inference to the best explanation have to say, in summary, for the attempt to formulate and evaluate a theistic position in discussion with the sciences? Based on the foregoing, I suggest the following four-step method:

1. Specification of a particular explanatory question and domain.
2. Initial formulation of the theistic hypothesis as a potential explanation for the data in question.
3. Identification of foils—that is, plausible competing explanations of the same data—leading to the formulation of a short list of likely candidates (in some cases, the theist's hardest job may be to show that her view has sufficient initial plausibility for it to be included at all as a potential explanation in the explanatory competition).
4. Drawing of an inference to the best explanation (an explanatory hypothesis), according to the method and criteria sketched above.

Step 1 expresses my contention at the beginning of this paper that the framework of explanations, rather than that of the theistic proofs or probability calculations, is the manner in which theistic claims are most appropriately advanced and defended. The possibility of achieving 2 in any sort of detail clearly depends on 1: only when the theist is addressing a well-defined explanatory task will she know how to develop a plausible and appropriate theistic hypothesis.[7] Theistic beliefs achieve their "purchase" through contact with the evolution debate or with cosmological theories or with discussions of human nature and values. Step 3 demonstrates a commitment to a fair and critical competition between serious contenders by making clear to all what exactly the range of options is. If any obvious candidate has been excluded, it becomes easy to recognize and redress the omission. Only when these first three steps have been undertaken can one turn to 4 with the hope of reaching a rational decision regarding the best available explanation.

With these steps before us, it now becomes possible to recognize several widespread fallacies:

- The failure of natural theology in both its deductive and its strictly inductive guises does not mean the end of the rationality of theistic claims. "Either deductive or inductive" represents a false dichotomy.

- Equally false is the dichotomy "either subjective or objective." The idea that an omniscient neutral observer could reach a single objective decision regarding each explanatory competition is a myth. Differences of explanatory interest mandate an irreducibly contextual element in explanation. Nonetheless, it does not follow that all explanations are created equal. We are often able to agree on a common body of evidence that needs to be explained, on the inadequacy of certain (past or present) explanatory proposals, and on the standards or "explanatory virtues" to which potential explanations should be subjected. Our neutral observer may not be able to legislate every argument, even in principle, but she *could* be in the position to recognize which is the best explanation relative to a particular set of explanatory interests, a specific set of background assumptions and theoretical beliefs, and a given body of evidence. In many cases, I suggest, the shared context is already clearly in place, thanks to developments in science or to certain common features of the human condition.
- In order to assert the (provisional) truth of an explanatory hypothesis, it is not necessary to show that all possible explanations have been canvassed or that no better explanation could ever be found. One is responsible only for the plausible explanations (according to our best lights). To put it differently: explanation involves a comparative, not an absolute, judgment.

This theory of explanation reflects a more general paradigm shift regarding the rationality of both scientific and meta-physical debates, a shift that we cannot fully explore here. In place of the inductive epistemologies of the logical positivists or the strict falsificationism of Karl Popper, it presupposes the notion of scientific and theological "research programs."[8] Likewise, in place of foundationalist understandings of knowledge it presupposes a coherentialist framework.[9] This brings inference to the best explanation into close contact with the "holistic view" of scientific explanation recently proposed by Philip Kitcher:

> [The holistic view] holds that scientific understanding increases as we decrease the number of independent assumptions that are required to explain what goes on in the world. It seeks laws and principles of the utmost generality and depth. . . . Explanations serve to organize and systematize our knowledge in the most efficient and coherent possible fashion. Understanding, on this view, involves having a world-picture—a scientific *Weltanschauung*—and seeing how various aspects of the world and our experience of it fit into that picture.[10] (Kitcher 1989, 182)

Reduced to its core, we could say that inference to the best explanation reflects the sense that the theist's first obligation is not to marshall evidence and arguments in order to justify her introduction of the term *God*. God talk is *not* epistemically inadmissible until proven otherwise. Rather, as Robert Prevost (1990, 159) has suggested, "the concept of

God is introduced as a theoretical entity postulated to explain various phenomena." Once she has defined a relevant explanatory project, the theist's job is to explicate the theistic hypothesis in sufficient detail and with sufficient rigor that its explanatory potential (or lack of same!) can be critically assessed.

IV

Although I have drawn freely from various theorists of inference to the best explanation, I have not paused to criticize the details of their work. One could ask, for example, whether a certain positivism doesn't characterize portions of Lipton's theory—as, for example, in his stress on causal *mechanisms* or his preference for a reduction of teleological explanations to histories of the interaction of efficient causes (for example, 1991, 132); one also wonders whether our "philosophical intuitions" are as sure a guide to truth as Lipton seems to take them (1991, 183 f.). Likewise, I believe Prevost's reservations (1990, 108 f.) against an IBE analysis of theism in general, as opposed to specifically Christian theism, are mistaken. Surely there are knotty problems with establishing connections between scientific theories and the highly specific historical and doctrinal claims of Christianity, problems that do not arise in the same way for a more minimal personalist theism. Still, the goal of these pages has been less to criticize others than to appropriate the portions of inference to the best explanation that are crucial for discussions with science—as well as for theology, metaphysics, and all other disciplines that make explanatory claims.

I have argued on the one hand that inference to the best explanation removes a certain burden from theistic language, the burden of being considered guilty until proven otherwise. On the other hand, I must emphasize that IBE also imposes some burdens of its own. Where serious explanations exist in some scientific field, we are not free to dismiss them for theological reasons without ourselves entering the explanatory fray and offering commensurable theistic explanations of our own that are as good as or better than the other options. I can put this more strongly: merely criticizing is not enough; we must also show the explanatory strength of our own theories. Consider, for example, the much-discussed attack on Darwinism by the Christian law professor Phillip Johnson, *Darwin on Trial* (1991). Johnson points out numerous difficulties with the Darwinist theory, including one or two serious ones. According to inference to the best explanation, the question for Darwinism (or for any scientific theory) is not "Is it evidentially and inferentially infallible?" but, rather, "Is it the best available explanation from among the competitors?" To decide *this* question, Johnson must do what he has not yet done: he must specify his Christian alternative to Darwinian explanations. Should

he be able to outline a stronger competing explanation, the intellectual community should indeed reject Darwinism; until then, it is under no obligation to do so. Alternatively, Johnson could use the foils technique to change the emphasis of the explanatory question without losing contact with the biological data. For example, instead of accepting the Darwinist's question Why do the mechanisms of genetic variation and natural selection, *rather than some other set of naturalistic mechanisms*, better explain the development of biological life forms? he might ask, Which theory—that of Darwinian (nonteleological) evolution or that of guided (directed) evolution—better accounts for the development of life, including the problem of the major transitions (nonliving structures to living cells, fundamental phyla, the appearance of *Homo sapiens*)? Given a more precise specification of the relevant foils, one might be able to expand the set of plausible options, perhaps leading to the result that Johnson is looking for—the explanatory superiority of introducing divine guidance of the evolutionary process.[11]

To say that the theist must "enter the fray" with the scientist is not, of course, to hold that she must begin with a full-fledged Christian biology. At first we can be agnostic about the details of theistic evolution, arguing only that, given the choice between theist and purely naturalistic explanations, there are good reasons for preferring the theistic side. As the debate continues, however, it becomes incumbent upon theists to provide a full explanatory account, and this means having more to say about the how question. At *this* point we must spell out the degree to which God may have used natural-evolutionary mechanisms and the reasons that God may have done so (for example, the importance of preserving the integrity of the created world, the fact that miraculous interventions in history are better understood as serving salvific purposes, the unanswerability of the problem of evil if one posits repeated interventions of God in the world, and so forth).[12]

We may agree, in closing, with Lipton (1991, 130): "We are members of a species obsessed with making inferences and giving explanations" (although one wants to ask him, But what is the best explanation of *this* fact?!). The importance of assessing theistic claims in terms of their success as explanations cannot be overestimated. I have not claimed that inference to the best explanation eliminates every incommensurability between science and religion or between competing philosophical proposals. For example, the theologian is finally interested in the quest for the ultimate explanation of the universe (Clayton 1989, 128–32), whereas the scientist rightly focuses on quantitative explanations and the "mechanism" requirement (Lipton 1991, 118). IBE, however, does draw attention to our common interest in rationality and to some common standards. As Ian Barbour

notes, although "religion cannot claim to be scientific or to be able to conform to the standards of science . . . it can exemplify some of the same spirit of inquiry found in science" (1990, 89f.). That is, it *is* possible to maximize the points of contact between religion and science, and the theory of inference to the best explanation provides the best methodological basis for doing so.

NOTES

Significant segments of the argument presented here were developed during a series of discussions with Steven Knapp, and some formulations are his. The author is grateful also to Alvin Plantinga for an extended correspondence on this topic, to Richard Paul and Lorenz Puntel, and to the Alexander von Humboldt Foundation for a grant supporting this research.

1. The method of inference to the best explanation has its classic expression in the work of Gilbert Harmann and can be traced back to the "method of abduction" in the writings of Charles S. Peirce. In recent literature, however, it has found a new application in the context of post-Gettier epistemology and postformalist, post-Kuhnian philosophy of science.
2. As Robert Prevost (1990, 2) puts it: "Explanatory arguments allow for degrees of support for alternative explanations. An explanation may be a poor one or a good one; yet in both cases the explanation is supported by the evidence, even if only minimally."
3. The best-known contemporary advocate of a probabilistic natural theology is Richard Swinburne. See his *An Introduction to Confirmation Theory* (1973), for example, and its application in *The Existence of God* (1979).
4. Lipton's book is as significant within the philosophy of science as it is in broader applications such as the present one. The following section draws heavily from his own exposition.
5. On the contrastive nature of explanation, see also Alan Garfinkel (1981, 28–41) and Bas van Fraassen (1980, 126–29).
6. Bengt Hannson, "Explanations—Of What?" cited in van Fraassen (1977, 143–150).
7. This is not to say that there is no place for an "internal" explication of (for example) Christian beliefs, as in traditional Christian theology; indeed, one could argue that the subsequent steps presuppose that one has already carried out this sort of internal work. I presuppose here that one is *also* interested in bringing one's religious beliefs into contact with the results of science or philosophy.
8. See the works by Clayton (1989) and Nancey Murphy (1990). There are many parallels between inference to the best explanation and the research-program methodology; for instance, both reject a direct movement from evidence to theory, both allow for the effect of contextual or pragmatic factors in theorizing; and both require comparative judgments between existing options. I suggest that the research-program approach, as a theory of (scientific) rationality, represents the more general theory, whereas inference to the best explanation concentrates on the specific nature and structure of explanations within the context of this more general theory of rationality.
9. Out of the huge bank of literature on this subject, see especially Nicholas Rescher (1973), Laurence Bonjour (1985), and Lorenz Puntel (1991).
10. Lipton (1991, 68) speaks similarly of an "explanatory detour," which in science "often requires 'vertical' inference to explanations in terms of unobserved and often unobservable entities and processes, and Inference to the Best Explanation seems particularly well equipped to account for this process." From a theological perspective it is certainly interesting to note that both Kitcher and Lipton deny any rational disadvantage to explanations that appeal to unobservable entities. Introducing such theoretical entities—and God is certainly one such!—does not require an a priori rational justification. Rather, each explanation is justified *after its introduction* by means of its explanatory success (or rejected because of explanatory impotence).
11. Of course, Johnson might respond (as Alvin Plantinga did in correspondence) that, if an explanation is sufficiently weak, it should not be believed, whether or not one has a better explanation to put in its place. This is defensible: sometimes the only existing explanation is a bad one (cf. Lipton 1991, 172, 177). Three factors mitigate against Johnson's (and Plantinga's) use of this observation in the debate with Darwinism, however. First, inference to the best explanation is a fundamentally comparative endeavor. In the typical case, my rejection of one explanation will be more convincing if I am able to propose an alternative. Second, inference to the best explanation requires one to specify the explanatory question and the set of plausible explanations *before* making

final conclusions for or against particular proposals. It is in practice rare to carry out this exercise without encountering (or being able to propose!) at least a small set of viable alternatives. Finally, the foils mechanism allows one to specify, and hence to modify, the precise explanatory question to be answered. If I cannot tell you precisely whose watch we have discovered in the forest—and, hence, cannot falsify your claim that it belongs to Saddam Hussein—I can nonetheless argue (convincingly, I think) that it is more likely to belong to an average citizen than to a head of state. If I disagree with an explanation and think that my disagreement is rational, I am obligated to provide an alternative *at the relevant level of detail*—and it is not difficult to imagine how one would engage in a rational discussion with a conversation partner about what *is* the appropriate level of detail for the best explanation of a given explanandum.

12. It would seem, for example, that a Christian view would include the following claims: (1) that God had a direct causal role in the creation of life, including the major "kinds"; (2) that biological mechanisms control much of the development (such as branching) within kinds; and (3) that the emergence of human life, at least, requires one to posit a divine guidance of the process. Regarding claim 3 we might hold either (*a*) that God created not only human beings but each of the various biological kinds directly, either at the outset or in subsequent interventions, or (*b*) that God's intervention in the case of human beings was qualitatively different from God's intervention in other cases. (It is interesting to note that in conversation Phillip Johnson admitted the need for developing a "positive" program of this sort.)

References

Barbour, Ian. 1990. *Religion in an Age of Science: The Gifford Lectures 1989–1991, Vol. One*. San Francisco: HarperCollins.

Bonjour, Laurence. 1985. *The Structure of Empirical Knowledge*. Cambridge: Harvard Univ. Press.

Clayton, Philip. 1989. *Explanation from Physics to Philosophy: An Essay in Rationality and Religion*. New Haven: Yale Univ. Press.

_____. 1996. *Das Gottesproblem. Gott und Unendlichkeit in der neuzeitlichen Philosophie*. Paderborn: Ferdinand Schöningh Verlag. Forthcoming in English as *Toward a Pluralistic Metaphysics: Models of God in Early Modern Philosophy.*

Gallie, W. B. 1955–56. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* 56: 121–46.

Garfinkel, Alan. 1981. *Forms of Explanation*. New Haven: Yale Univ. Press.

Harmann, Gilbert. 1965. "The Inference to the Best Explanation." *Philosophical Review* 74: 88–95.

Hick, John, and Michael Goulder. 1983. *Why Believe in God?* London: SCM Press.

Kitcher, Philip, and Wesley Salmon, eds. 1989. *Scientific Explanation*. Minnesota Studies in the Philosophy of Science, vol. 13. Minneapolis: Univ. of Minnesota Press.

Johnson, Phillip E. 1991. *Darwin on Trial*. Washington, D.C.: Regnery Gateway.

Lipton, Peter. 1991. *Inference to the Best Explanation*. London: Routledge and Kegan Paul.

Murphy, Nancey C. 1990. *Theology in the Age of Scientific Reasoning*. Ithaca, N. Y.: Cornell Univ. Press.

Prevost, Robert. 1990. *Probability and Theistic Explanation*. Oxford: Clarendon.

Puntel, Lorenz. 1991. *Grundlagen einer Theorie der Wahrheit*. Berlin: de Gruyter.

Rescher, Nicholas. 1973. *The Coherence Theory of Truth*. Oxford: Clarendon.

Swinburne, Richard. 1973. *An Introduction to Confirmation Theory.* London: Methuen.

_____. 1979. *The Existence of God*. Oxford: Clarendon Press.

van Fraassen, Bas. 1977. "The Pragmatics of Explanation." *American Philosophical Quarterly* 14: 143–150.

_____. 1980. *The Scientific Image*. Oxford: Oxford Univ. Press.