Bayesian Confirmation of Theories That Incorporate Idealizations*

Michael J. Shaffer†‡
The University of North Carolina at Wilmington

Following Nancy Cartwright and others, I suggest that most (if not all) theories incorporate, or depend on, one or more idealizing assumptions. I then argue that such theories ought to be regimented as counterfactuals, the antecedents of which are simplifying assumptions. If this account of the logical form of theories is granted, then a serious problem arises for Bayesians concerning the prior probabilities of theories that have counterfactual form. If no such probabilities can be assigned, then posterior probabilities will be undefined, as the latter are defined in terms of the former. I argue here that the most plausible attempts to address the problem of probabilities of conditionals fail to help Bayesians, and, hence, that Bayesians are faced with a new problem. In so far as these proposed solutions fail, I argue that Bayesians must give up Bayesianism or accept the counterintuitive view that no theories that incorporate any idealizations have ever really been confirmed to any extent whatsoever. Moreover, as it appears that the latter horn of this dilemma is highly implausible, we are left with the conclusion that Bayesianism should be rejected, at least as it stands.

1. Introduction. In this paper I present a new criticism of (subjective) Bayesian confirmation theory based on the empirically grounded observation that most, if not all, theories hold true only in models that incorporate, or depend on, one or more idealizing conditions. I will argue that

*Received December 1999, revised October 2000.

†Send requests for reprints to the author, Department of Philosophy and Religion, University of North Carolina at Wilmington, Wilmington, NC 28403–3297. shafferm@uncwil.edu.

‡I would like to thank Harold Brown, Risto Hilpinen, A.J. Kreider, Pawel Kawalec, two anonymous referees, and the editors for comments on earlier drafts of this paper. I would also like to thank the University of Miami Philosophy Department for providing me with support while this work was done. Earlier versions of this paper were presented at the Second Summer School for the Theory of Knowledge held in Warsaw in 1999 and at the 1999 meeting of the Florida Philosophical Association in Miami. I thank the participants of both conferences for their comments.

Philosophy of Science, 68 (March 2001) pp. 36–52. 0031-8248/2001/6801-0003$2.00

if this assertion about the ubiquity of idealization in the sciences is true, then Bayesian confirmation theory does not, at least as it stands, provide us with a coherent measure of the support that most, if not all, scientific theories gain from empirical evidence. This is due to the peculiar logical form of such theories. I will refer to this problem as the Bayesian problem of idealization. Subsequent to the presentation of this criticism, two possible solutions to the Bayesian problem of idealization will be considered, and I will argue that these proposals are unsuccessful. In virtue of these results I will suggest that Bayesianism ought to be given up, at least with respect to those types of theories discussed herein.

2. Idealization in the Sciences. That scientists frequently employ the operation of idealization is easily seen in any serious empirical examination of virtually any area of science. Numerous examples can be found in physics, biology and economics, as well as the other special sciences, and this has been true at least since Galileo made that method more or less legitimate.[1] However, the extent to which the sciences employ the operation of idealization has often been overlooked or, at least understated, in the philosophy of science. To some extent this myopia on the part of philosophers of science has recently been cured. In particular, the works of Nancy Cartwright, Ian Hacking, R. I. G. Hughes, Ronald Laymon, Leszeck Nowak, and Ernan McMullin have addressed issues concerning the ubiquity and function of idealizations in the sciences.[2]

For the most part, though, little or no attention has been paid to the formal issues that arise concerning both the logical properties of idealizing assumptions and the manner in which claims that depend on idealizing assumptions are empirically confirmed. In what follows I will address aspects of both of these issues. In addressing the issue of confirmation of such theories, I will present a novel criticism of Bayesian confirmation theories of the subjective sort, grounded in the claim I will make about the logical form of theories that depend on idealizing assumptions. Although I will not offer a full presentation of the logic of such expressions here, I believe that there are good reasons to accept the basic regimentation of such expressions proposed in Sections 3.1 and 3.2.[3] In any case, from a purely pragmatic perspective, I believe that the account of idealization considered here provides the beginnings of a much-needed formal

1. For various perspectives on Galileo's methods and their subsequent impact see Butts and Pitt 1978, Pitt 1992, Shapere 1974, McMullin 1985, and Koyré 1968.

2. See, for example, Nowak 1980, Cartwright 1983, McMullin 1985, Hughes 1990, Hacking 1983, Laymon 1989 and various other works by these philosophers.

3. For a considerably more detailed presentation and justification of the logic of idealization see Shaffer 2000. Therein the logic of idealization is referred to as VCP.

grounding for the claims made by those philosophers of science concerned with the role of idealization in the sciences noted above.

3. The Function of Idealizations and Their Logic. Idealization is the operation of counterfactual simplification.[4] It seems obvious that this is so when we examine particular historical cases of the use of idealization. Essentially what is being done in employing idealization is that a particular theoretical claim is being asserted on the basis of one or more simplifying assumptions that are known to be false. In many cases these simplifying assumptions are made explicit, but this is clearly not always the case, and so, often times, one has to do some digging to uncover the simplifying assumptions that are lurking behind a given theoretical claim. However, in uncovering the logical form of theories and ascertaining the meanings of theories it is absolutely crucial that we recognize and acknowledge the idealizing assumptions that are explicitly or implicitly presupposed within the context of any such theory.

When one starts looking at even the simplest cases in detail it quickly becomes apparent that idealizations are ubiquitous in all of the sciences, and I submit that this is so because the world is a highly complex place and our computational and descriptive powers are subject to numerous physical limitations.[5] So, in general, it seems to be the case that we construct simplified models of complex physical processes in order to secure computational tractability by simplifying the equations that describe a given type of situation.[6] Ultimately, this procedure appears to be employed in the sciences as a matter of physical necessity, and, again, this is because there are various physical/computational constraints imposed on both our cognitive abilities and those of computers.[7]

4. It is important to distinguish here between the concepts of one model being a simplification of another and one model being simpler than another. The former notion involves the concept of one model contracted by removing elements of some base model in question, while the latter is more plausibly construed as some absolute measure of the overall simplicity of models. This distinction is crucial for the semantics of VCP developed in Shaffer 2000.

5. I also tend to agree with the Humean observation, made in the context of our inductive practices, that humans have a natural tendency to regard the world as being much more orderly and homogenous than it probably is. Making this unwarranted assumption simply makes things easier to deal with, and this point is echoed in much of the work in artificial intelligence concerning the frame problem. For discussion of this problem see Pylyshyn 1986 and Shoham 1988. I believe that these sorts of observations concerning our tendency to simplify things apply equally well in the context of idealization.

6. The same point is made in Redhead 1980.

7. Detailed considerations of the physical limitations on computability can be found in Leff and Rex 1990, Casti and Karlqvist 1996, Geroch and Hartle 1986, and Pitowsky 1990. Details concerning mathematical aspects of computationally intractable problems

3.1 Analytical Mechanics of Projectile Motion: An Analysis. To illustrate how the procedure of idealization is employed in practice, consider the following example from Arthur and Fenster's *Mechanics* (1969, Ch. 7). They present an account of the general motion of particles. As one example of these sorts of motions they present the classical example of projectile motions. They offer three progressively more complex accounts of such motions, and the presentation of these accounts is qualified explicitly with the following caveat:

> In studying the motion of projectiles, we begin with a much simplified case. As the original assumptions are changed to improve the approximation of the "real case," the equations become increasingly complex. In practice, a point is eventually reached in which numerical techniques suitable for computer solution are employed. The reader must be aware that approximations and simplifications limit the applicability of the results. (Arthur and Fenster 1969, 235)

In an exemplary presentation, this warning is explained in that they actually list the following idealizing assumptions that they are making in their first analysis of projectile motion:

a. *The projectile is a point mass or particle.* In a more accurate analysis, it would be considered a body possessing finite volume and a definite surface configuration. Our concern would then be with the motion of the mass center. The *attitude* of the projectile, described by the angles between reference axes in the projectile and a convenient external coordinate reference, is related to the air drag and would therefore enter into the formulation of this problem.
b. *The earth is nonrotating.* If greater accuracy is required, the accelerated or noninertial motion of the earth beneath the projectile must be taken into account. In this chapter the earth is used as a reference for which Newton's laws are assumed valid.
c. *The gravitational field is constant and acts perpendicular to the surface of a flat earth.* For distances small in comparison with the earth's radius, the flat-earth assumption yields good results.
d. *The air offers no resistance to motion; that is, motion occurs as it would in a vacuum.* Actually, air friction is important. It depends upon projectile attitude, wind velocity, air density, air viscosity, projectile configuration, and projectile speed.
e. *Motion occurs in a plane.* (Arthur and Fenster 1969, 235–236)

---

are given extensive consideration in Garey and Johnson 1979. Considerations about the limitations on human computational abilities are discussed in Stich 1990, Brown 1988, and Cherniak 1986.

So, even something as basic as the mechanical account of projectile motion is presented in a way that depends on numerous, more or less serious, simplifying assumptions, and, recalling the warning quoted earlier, it is clear that the simpler accounts are introduced due to their computational tractability. Assuming a., b., c., d., and e. allows us to analytically solve the simpler sets of differential equations, whereas eliminating one or more of those assumptions makes those computations considerably more difficult, and in some cases analytically unsolvable.

The first account of projectile motion Arthur and Fenster provide, projectile analysis 1 (hereafter, PA1), holds true only under all of a., b., c., d., and e.. They derive the following set of equations to describe the motion of a projectile fired from a point, in terms of the components of its motion in the X (horizontal) and Y (vertical) directions:

(PA1.1) $x = v_{ox}t = (v_0\cos\alpha)t,$
(PA1.2) $\partial x/\partial t = v_x = v_o\cos\alpha_0,$
(PA1.3) $y = -1/2gt^2 + (v_0\sin\alpha_0)t,$
(PA1.4) $\partial y/\partial t = v_y = -gt + v_0\sin\alpha_0,$

where v is velocity, t is time, α is the angle of the initial velocity relative to the flat earth, and g is the gravitational constant. So this example can be logically analyzed as asserting:

(I1) Were it the case that a., b., c., d., and e.,
then it would be the case that (PA1.1)–(PA1.4),

where a., b., c., d., and e. are known to be false, but (PA1.1)–(PA1.4) are easily solved for some given set of initial conditions.

Following the presentation of (PA1), Arthur and Fenster introduce a second, considerably more complicated, account of projectile motion (PA2). In (PA2) the idealizing assumption d. is eliminated, and so they take into account fluid resistance in the medium in which the motion occurs. In eliminating d. the consequent of (I1) must be suitably modified so as to incorporate two types of frictional forces. The first force, lift, is perpendicular to the relative velocity of the approach of the fluid and the object, and the second force, drag, is parallel to the velocity of the approach of the fluid and the object. The drag and lift coefficients are as follows:

$$C_D = F_D/(1/2\rho u^2_\infty A),$$
$$C_L = F_L/(1/2\rho u^2_\infty A),$$

where F_D and F_L are the drag and lift forces, ρ is the mass density of the fluid, u_∞ is the velocity of the undisturbed fluid relative to the object, and

A is the is the area of the object projected on a plane perpendicular to the undisturbed fluid velocity. Each of these forces is described by a dimensionless coefficient, and in their presentation of how we are to incorporate frictional forces into (PA1), Arthur and Fenster assume that the projectile is spherical so that the coefficient of lift is 0.

Having introduced this further stipulation, they continue and note that C_D, the coefficient of drag, is often related to another dimensionless variable, the Reynolds number. The Reynolds number is defined as follows:

$$\mathrm{Re} = \rho u_\infty L/\mu.$$

L is the characteristic length of an object, μ is the viscosity of the fluid, and the other terms are as before. In any case, Arthur and Fenster explain that when the Reynolds number is low, the drag force is (approximately) proportional to the projectile velocity and one can derive the correlates of (PA1.1)–(PA1.4) for the frictional case. They are as follows:

(PA2.1) $x = -(m/b)v_{0x}e^{-(b/m)t} + C_2,$
(PA2.2) $v_x = v_{0x}e^{-(b/m)t},$
(PA2.3) $y = -(b/m)(v_{0y} + (m/b)g)e^{-(b/m)t} - mgt/b + C_4,$
(PA2.4) $v_y = (v_{0}y + mg/b)e^{-(b/m)t} - mg/b,$

where m is mass, b is the linear drag coefficient, and C_2 and C_4 are determined by initial conditions. So, this theoretical claim should be construed as follows:

(I2) If a., b., c., and e. were the case,
then (PA2.1)–(PA2.4) would be the case.

But, the equations that constitute (PA2) are, generally, much more difficult to solve analytically than those that constitute (PA1). Tractability is further reduced in the analysis of (PA3) where motions that involve higher Reynolds numbers are considered. As a result, it is often much simpler to accept (PA1) to secure computational tractability at the expense of realism, even if doing so results in greater disagreement with the observed evidence, and this is done with full awareness that (I2) is a much better explanation of that evidence than (I1).

3.2 The Logical Form of Idealized Theories. Regardless of whether or not presuppositions like those made in the cases of (PA1) and (PA2) are explicitly stated, once we recognize the function of idealizing assumptions the logical form of such idealizations should be clear. When we claim that a theory holds only in some idealized model, or under some idealizing conditions, we are claiming that that theory is true only on the basis of one or more *counterfactual* simplifying assumptions or conditions. As such, the logical form of a theory that incorporates one or more idealizing assumptions is that of a counterfactual conditional, i.e. $A > C$.[8] But, whereas standard counterfactual conditionals are about how things are in other close complete possible worlds similar to a given world, I want to suggest that idealizing counterfactuals, insofar as idealizing involves simplification, are about how things are in close simpler, incomplete or partial, possible worlds similar to a given world.[9]

So, the logic of idealization is a conditional logic, and I find it to be rather strange that this has gone unnoticed or ignored by those philosophers of science concerned with the role of idealization in the sciences.[10] However, the details concerning which conditional logic most appropriately represents the logic of such expressions can be ignored in what follows, as it in no way affects the problem that I will raise. Given this brief examination of the logical form of theories that incorporate idealizations, we can now turn our attention to the treatment of theories like (I1) and (I2) by Bayesian confirmation theory.

4. The Basics of Bayesian Confirmation Theory. To generate the problem for Bayesian confirmation theory to be presented in Section 4.1, all that is necessary is that I be granted the assumption that theories incorporating idealizing conditions ought to be construed as counterfactuals, and I believe that this thesis is not open to question. However, before this criticism of Bayesian confirmation theory is presented explicitly I will outline the basic principles of Bayesianism for the benefit of those who might not be particularly familiar with that view.

8. As far as I have been able to discover, this point has been explicitly noted only in Hanson 1965 and Dalla Chiara 1992. Hanson recognized that no actual systems satisfy the presuppositions of Newton's first law, and, given the philosophical fashions of the time, concluded that it was unfalsifiable. Dalla Chiara, on the other hand, briefly comments that the boundary conditions assumed in the contexts of theories are not usually satisfied in actuality.

However, related methodological issues concerning the completeness of theories and partial explanations are discussed in Fetzer and Nute 1979, Fetzer 1981, and Railton 1981. Of special interest in this regard is Fetzer's criterion of maximal specificity. This criterion, essentially, states that theories or laws can be true only when all factors that make a difference to the phenomenon being described are taken into account. Of course, it is unlikely that such a criterion is typically satisfied, if it is even ever satisfied, when dealing with actual theories.

9. In this analysis I follow Tore Langholm's suggestion from Langholm 1996.

10. For a more detailed investigation of the logic of idealization I refer the reader to Shaffer 2000. In this work the logic of idealization is construed as an extension of Lewis' conditional logic VC that allows models to be more or less complete. As such, VCP is both three valued and nonmonotonic and it shares many features in common with default logics and circumscription logics. For some detailed considerations about logics of this sort see Brewka, Dix, and Konolige 1997, Makinson 1994, Poole 1994, and Lifschitz 1994.

Bayesianism is the most well entrenched theory of confirmation. Standard Bayesian confirmation theory holds that degrees of belief ought to conform to the axioms of the probability calculus. This requirement is referred to as the requirement of coherence, and is typically supported by appeal to various forms of so called Dutch book arguments.[11] These arguments are designed to show that it would be irrational to have a probability distribution over one's beliefs that did not obey the probability calculus.

In any case, following the presentation in Howson and Urbach 1993, the axioms of the probability calculus are as follows:

(A.1) $P(a) \geq 0$ for all a in the domain of $P(\cdot)$.
(A.2) $P(t) = 1$ if t is a tautology.
(A.3) $P(a \vee b) = P(a) + P(b)$ if a and b and $a \vee b$ are all in the domain of $P(\cdot)$, and a and b are mutually exclusive.

In (A.1)–(A.3), $P(\cdot)$ is a function whose domain is, normally, a complete set of statements closed under Boolean operations.[12] Furthermore, from these axioms Bayes' Theorem can be derived, and Bayes' theorem, in its most well known form, holds that,

(B.1) $P(h \mid e) = P(e \mid h)P(h)/P(e)$, provided that $P(e) > 0$.

In other words, (B.1) tells us how well e supports h. The sort of Bayesian confirmation with which I am concerned interprets the function $P(\cdot)$ as the assignments of credal, or subjective, probabilities over a given agent's set of beliefs governed only by the axioms of the probability calculus.[13] Given this brief exposition of standard Bayesian methodology, we can now ask how theories that incorporate idealizations might be regarded from the Bayesian perspective.

4.1. A New Problem for Bayesian Confirmation Theory. Consider how we would substitute theories of the form $A > C$, such as the theory of projectile motion (I1) described in Section 3.1, into Bayes' theorem. Recall that Bayes' theorem says that the posterior probability of an hypothesis

11. See Earman 1992 and Rosenkrantz 1981 for details concerning Dutch book arguments.

12. Things are a bit more difficult when probabilities are introduced into the context of VCP, as the standard Boolean connectives are replaced in VCP by the strong Kleene connectives, but for the sake of simplicity I will ignore this issue in what follows. See the references in footnotes 9 and 14 for explanations of the genesis of this problem. Also, it is useful to compare and contrast VCP with the conditional event algebra developed in Walker 1994.

13. For considerations of other interpretations of the probability calculus I refer the reader to Good 1983 and Howson 1995.

conditional on the evidence is equal to the product of the probability of the evidence conditional on the hypothesis, the likelihood, and the prior probability of the hypothesis, divided by the probability of the evidence, provided the probability of the evidence does not equal zero. Substituting $A > C$ for h we get,

(B.2.) $P(A > C \mid e) = P(e \mid A > C)P(A > C)/P(e)$, provided $P(e) > 0$.

The problem I wish to point out concerns the nature of subjective prior probabilities of theories with the form of counterfactuals. Standard criticisms of subjective Bayesianism focus on the subjective character of prior probabilities (see especially Glymour 1981, Kyburg 1978, Brown 1994, and Salmon 1990). While I tend to be sympathetic to these criticisms, I will focus here on a rather different problem, viz., how are we to understand subjective prior probabilities of conditionals like $P(A > C)$ in B.2.?

If Bayes' theorem is to be well defined in a given application, it must be the case that we can assign some meaningful value to each of the terms in that expression. But let us consider how we might assign a subjective prior probability to a theory that holds only under one or more counterfactual idealizing conditions. In essence, we are asking ourselves to assign a probability to an expression about what would be the case in close possible worlds that are similar to a given world, but simplified in some respect(s). Initially one might not really think that there is a problem here at all, but this is by no means the case.

The most plausible suggestion concerning how the probabilities of conditionals ought to be construed is that the probability of a conditional should be interpreted as the conditional probability of the consequent given the antecedent.[14]

$P(A > C) = P(C \mid A)$ for all A, C in the domain of P with $P(A) > 0$,

and

$P(C \mid A) = P(CA)/P(A)$ provided $P(A) \neq 0$.

Alan Hàjek (1989) has proposed the acronym 'C.C.C.P.' to refer to this account (the conditional construal of conditional probability), and I shall follow this convention throughout.

Unfortunately for the Bayesian, as Davis Lewis and others have demonstrated, C.C.C.P. cannot be correct on pain of triviality. Based on some rather minimal assumptions, Lewis (1976) showed that any language having a universal probability conditional is a trivial language, and, hence, by reductio C.C.C.P. must be rejected.[15] Furthermore, in Hájek 1989,

14. A detailed and illuminating history of attempts to interpret probabilities of conditionals is presented in Milne 1997.

15. For specific details I refer the reader to Lewis 1976, Lewis 1986, and McGee 1989.

C.C.C.P. was proved to be trivial under considerably weaker assumptions than those originally made in Lewis 1976.

For the Bayesian, this result becomes problematic with respect to the sorts of theoretical claims described above, such as (I1) and (I2). In point of fact, if one agrees with the basic point about the ubiquity of idealizations raised by Cartwright (1983), Hacking (1983), McMullin (1985), Nowak (1980), and others and with the account of the logical form of theories that incorporate idealizations given above, then most, if not all, theoretical claims made in the course of scientific activity cannot be confirmed in the way Bayesians claim. To drive the point home, this is because most, if not all, theoretical claims depend on idealizing assumptions, and theories that incorporate such idealizations ought to be construed logically as special sorts of counterfactual conditionals. If this is so and there is no extant suggestion for how to assign prior probabilities to counterfactuals, then the posterior probability of virtually every theoretical claim will be undefined in terms of Bayes' theorem.

It also seems intuitively obvious that given a set of experimental trials, say a set of actual projectile motions, (I2) should be better confirmed by that evidence than (I1) even though (I2) is not strictly speaking a correct explanation of that body of evidence; it omits several causal components of actual projectile motions. But, if all such theories lack well defined probabilities in terms of Bayes' theorem, then there is no way to compare the confirmational status of (I2) relative to (I1) on that body of evidence, and, hence, there will be no way to explain the common intuition that favors (I2) over (I1), given some fixed body of evidence.

In any case, this situation is unfortunate for the Bayesian as there does not seem to be any extant, coherent, suggestion as to how we are to non-trivially assign prior probabilities to indicative or counterfactual conditionals. This problem, *the Bayesian problem of idealization,* appears to have devastating consequences for Bayesianism. Unless Bayesians can come up with a coherent suggestion for how such probabilities are to be understood, either Bayesianism must be rejected or, given the ubiquity of idealizations, Bayesians must accept the rather counterintuitive conclusion that few, if any, scientific theories have ever been confirmed to any extent whatsoever.[16] Insofar as the latter alternative does not appear to be one that most Bayesians would be willing to accept, they must come up with some (non-trivial) account of how to understand probabilities of counterfactuals or they must give up Bayesiansim.

16. In effect, Bayesians would be committed to the view that only theories that do not incorporate idealizations can be confirmed in the manner suggested by Bayesians. There do not seem to be many theories of this sort.

5. Prospects for a Solution to the Bayesian Problem of Idealization. In response to Lewis' celebrated results, and the extensions thereof, two major proposals have arisen concerning the nature of conditionals and the probabilities of such expressions. First, Lewis has proposed a way for assigning probabilities to conditionals, called imaging, which will be considered in Section 5.1. Second, Isaac Levi (1996), Carlos Alchourrón, Peter Gärdenfors, and David Makinson (1985), and others have proposed various accounts that deny that conditionals are truth valued. Instead, they consider conditionals to be something like policies for belief revision; such policies have conditions of rational support, in lieu of truth conditions. Section 5.2 will consider this view. What we must now be concerned with here, first and foremost, is whether Bayesians can exploit these suggestions in order to solve the problem of idealization.

5.1. Lewis' Concept of Imaging. Subsequent to rejecting C.C.C.P. with respect to Stalnaker conditionals, as well as many other types of conditionals, Lewis (1976) suggested that probability conditionals should be understood as policies for *feigned* minimal belief revision, and that the probability of such a conditional should be understood to be the probability of the consequent, given the minimal revision of $P(\cdot)$ that makes the probability of the antecedent of the conditional equal to 1. Formally,

$$P(A > C) = P'(C), \text{ if } A \text{ is possible,}$$

where $P'(\cdot)$ is the minimally revised probability function that makes $P(A) = 1$. Lewis tells us that we are to understand this expression along the following lines. $P(\cdot)$ is to be understood as a function defined over a finite set of possible worlds, with each world having a probability $P(w)$. Furthermore, the probabilities defined on these worlds sum to 1, and the probability of a sentence, A for example, is the sum of the probabilities of the worlds where it is true. In this context the image on A of a given probability function is obtained by 'moving' the probability of each world over to the A-world closest to w.[17] Finally, the revision in question is supposed to be the minimal revision that makes A certain. In other words, the revision is to involve only those alterations necessary for making $P(A) = 1$.

But this analysis of the problem of the probabilities of conditionals does not seem to help the Bayesian in the least. Consider Bayes' theorem in light of this suggestion. Substituting $P'(C)$ for $P(A > C)$ yields

$$\text{(B.3)} \quad P(A > C \mid e) = P(e \mid A > C)\, P'(C)/P(e), \text{ provided } P(e) > 0.$$

What are we to make of this expression? What is the meaning of the product of the likelihood of the evidence given our theory and the prob-

17. Lewis also assumes that there is a unique closest A-world to w.

ability of the antecedent of our theory in terms of *some other probability function*, especially where the latter probability is the probability one *would* assign to the consequent after making the minimal revision of one's beliefs needed to make the probability of the antecedent equal to one. The two terms appear, in some sense, to be incommensurate, but, more importantly, there does not really seem to be any coherent way to assign a probability to $P'(C)$.[18] The revision in terms of which $P'(C)$ is defined *does not actually occur*, it is only a feigned revision. It only occurs counterfactually and it is not clear how in the world we are to assess what the value of $P'(C)$ should be. This is complicated by the fact that what counts as a minimal revision has not been satisfactorily fleshed out in the literature, and so, in any case, we appear to be at a loss to actually employ Lewis' solution in practice.[19]

Furthermore, Lewis' suggestion places us in a position that appears to involve a regress. In order to assess the numerical value associated with the image on A of $P(\cdot)$ we must accept another counterfactual concerning what we would believe if we were certain of A. This is because the belief revision is not an actual belief revision, and in order to accept this we would need to assign a probability to the counterfactual 'If I were certain of A (if it were the case that $P(A) = 1$), then my beliefs would be $\{\psi\}$', where $\{\psi\}$ is the set of my beliefs and probability ascriptions on those beliefs. Presumably, this counterfactual must be interpreted in terms of imaging as well, and so we must accept another counterfactual about that feigned revision, and so on.[20] As a result of these considerations, it does

18. The sense in which I suggest that they are incommensurate is that the revised probability function is not about what I believe at all, but about what I would believe and the likelihood is about what I believe. It is not all clear how the product of these probabilities is to be understood. A similar sort of problem might also be found in Jeffrey's conditionalization, a kinematical generalization of Bayesian conditionalization, where $P_{new}(A) = P_{old}(A|E)P_{new}(E) + P_{old}(A|{\sim}E)P_{new}|{\sim}E)$ for the pair $\{E, {\sim}E\}$. As in the case of (B.3), the products of these probabilities are rather difficult to interpret. It is similar to assuming that one can make sense of the product of the probability ascriptions of two different individuals, say $P_{John}(A|E)P_{James}(E)$. My partial beliefs at some time t_1 and at some subsequent time t_2 are like the beliefs of two different agents, and they are, potentially, ascribed over two different sample spaces. It does not seem to be obvious just what an expression involving products of these sorts of terms could mean.

19. See, for example, Gärdenfors 1984, Alchourrón, Gärdenfors, and Makinson 1985, and, especially Lindström and Rabinowicz 1989, Lindström and Rabinowicz 1990, and Lindström and Rabinowicz 1992.

20. In a bit more formal presentation this problem arises as follows. If $P(A > C) = P'(C)$ by imaging, then to assess the numerical value of $P'(C)$ an agent must accept the conditional $T(A) > \{B\}$ where $T(A)$ is the belief that a particular agent is certain that A and $\{B\}$ is that agent's set of beliefs and probability distribution over those beliefs. To accept $T(A) > \{B\}$ by Lewis' own admission is to assign a (high) probability to that sentence, so the agent must be able to evaluate $P(T(A) > \{B\})$ if the agent is to be able

not appear as if imaging will help the Bayesian avoid the problem of idealization, as imaging does not allow us to clearly specify a well defined prior probability for the kinds of theories we have been discussing.

5.2. The AGM/Levi Approach to Conditionals. In the spirit of F. P. Ramsey's and Ernest Adams' accounts of the nature of conditional expressions, some philosophers and computer scientists have adopted the view that conditional expressions do not have truth values (see Ramsey 1990; Adams 1975, 1976, 1993; Edgington 1986). Rather, they hold that conditionals ought to be regarded as various kinds of epistemic policies for belief revision. Although their views differ with respect to various details concerning the nature of such revisions, Isaac Levi, Carlos Alchourrón, Peter Gärdenfors, David Makinson, and others agree that conditionals of the sort we have been discussing should not be treated as assertions that have truth conditions (see Gärdenfors 1986, 1988; Alchourrón, Gärdenfors, and Makinson 1985; Arló Costa and Levi 1996; Levi 1996). Rather, they are to be treated as something like policies for updating or revising one's beliefs relative to what one already believes; in other words they are taken to be epistemic conditionals, and in lieu of truth conditions such conditionals have conditions for rational support relative to an antecedently given belief set.[21,22]

---

to assess $P(A > C)$. But, by imaging, $P(T(A) > \{B\}) = P''(\{B\})$, where $P''(\{B\})$ is the agent's beliefs and probability distribution on those beliefs were the agent certain that $T(A)$. Again, according to the definition of the concept of imaging this is only a feigned revision. So, in order to assign a numerical value to $P''(\{B\})$ the agent must accept a conditional about what that agent would believe if he were certain that he were certain that A, $T(T(A)) > \{B'\}$ (where $\{B'\}$ is that agent's suitably revised beliefs and his probability distribution on those beliefs). So, the agent must assign a numerical value to $P(T(T(A)) > \{B'\})$ where, by imaging $P(T(T(A)) > \{B'\}) = P'''(\{B'\})$, and so on ad infinitum. Moreover, there does not seem to be any obvious, non–ad hoc, way to stem this regress that results from the nature of imaging.

21. Lindström and Rabinowicz (1995) and Hansson (1995) make a more or less sharp distinction between epistemic and ontic conditionals. They hold that ontic conditionals have truth values while epistemic conditionals have conditions of rational support. They distinguish these two types of conditionals based on the idea that the latter are accepted relative to belief systems that do not need to be complete, whereas ontic conditionals are true or false only relative to complete models. In general I fail to see the difference between these two types of conditionals, and I believe that the distinction can be dissolved by allowing possible worlds to be partial (i.e., incomplete). However, I will ignore this issue for present purposes. For a more detailed discussion of partial worlds and conditionals see the essays in Doherty 1996.

22. The epistemic implications of this view are discussed at some length in Gärdenfors 1992. In this article Gärdenfors appears to ally the belief revision tradition with coherence theories of knowledge, and this provides some explanation of the AGM theorist's views concerning conditionals, at least qua their lack of truth values.

I will not dwell on this alternative at length, but simply note that, when applied to the kinds of conditionals discussed in Sections 3.1 and 3.2, this analysis is rather implausible and does not reflect scientific or everyday reasoning very well. As Stalnaker pointed out long ago, "many counterfactuals seem to be synthetic, and contingent, statements about unrealized possibilities" (1970, 42), and I take it that physicists who assert (I1) are making a substantive empirical claim about how things would behave under conditions that (contingently) do not occur in this world. It simply does not seem that such conditionals are merely about how beliefs might be revised. Rather, those conditionals are synthetic and contingent claims about how certain types of objects would behave in worlds that are simplifications of our world in some specified respect(s). In any case, the AGM/Levi analysis cannot be used to solve the Bayesian problem of idealization, because it does not allow for prior probabilities to be assigned to theories of the sort with which we are concerned. For if they do not have truth conditions, then they certainly cannot have probabilities of being true.

5.3. Provisional Conclusions and Prognoses. Given the failure of C.C.C.P., imaging, and the AGM/Levi account of conditionals to solve the Bayesian problem of idealization, Bayesians must apparently look elsewhere if they are to avoid the unpalatable conclusions that few, if any, theories have really ever been confirmed to any degree whatsoever, and that theories like (I2) are not better confirmed by actual evidence than simpler alternatives like (I1). There does, however, appear to be another way in which one might approach this new criticism of the Bayesian way, without accepting the counterintuitive consequences discussed above. One could simply give up Bayesianism (at least with respect to theories of the sorts described above) in favor of some other account of theory confirmation or acceptance, and perhaps this is what should be done.

My own intuitions tell me that both counterintuitive consequences could be avoided by appealing to inference to the best explanation (IBE), construed in terms of inference to the likeliest explanation. If our theory acceptance rule was something like:

(A) accept the theory that confers the greatest likelihood on the available evidence in virtue of our background knowledge,

then our accepting (I2) in lieu of (I1) would not involve subjective priors. Nor would it be incoherent to prefer (I2) to (I1), as the former is manifestly a better explanation of actual observed projectile motions, even if (I2) is not itself the perfect explanation of such motions.[23] These considerations

23. Including background knowledge in assessing the abductive strength of theories

and the criticisms raised above lead me to believe that the Peircean method of hypothesis, IBE, may be the only way to account for the acceptance of theories that involve idealizations, but, of course, much work needs to be done in defending this rather vague view.[24]

REFERENCES

Adams, Ernest (1975), *The Logic of Conditionals.* Dordrecht: D. Reidel.
——— (1976), "Prior Probabilities and Counterfactual Conditionals", in William L. Harper and Clifford A. Hooker (eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science,* Vol.1. Dordrecht: D. Reidel, 1–21.
——— (1993), "On the Rightness of Certain Counterfactuals", *Pacific Philosophical Quarterly* 74: 1–10.
Alchourrón, Carlos, Peter Gärdenfors, and David Makinson (1985), "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions", *The Journal of Symbolic Logic* 50: 510–530.
Arló Costa, Horacio and Isaac Levi (1996), "Two Notions of Epistemic Validity", *Synthese* 109: 217–262.
Arthur, Wallace and Saul Fenster (1969), *Mechanics.* New York: Holt, Rinehart, and Winston.
Brewka, Gerhard, Jurgen Dix, and Kurt Konolige (1997), *Nonmonotonic Reasoning: An Overview.* Stanford: CSLI.
Brown, Harold (1988), *Rationality.* New York: Routledge.
——— (1994), "Reason, Judgment and Bayes's Law", *Philosophy of Science* 61: 351–369.
Butts, Richard and Joseph Pitt (eds.) (1978), *New Perspectives on Galileo.* Dordrecht: D. Reidel.
Cartwright, Nancy (1983), *How the Laws of Physics Lie.* New York: Oxford University Press.
Casti, John L. and Anders Karlqvist (eds.) (1996), *Boundaries and Barriers: On the Limits to Scientific Knowledge.* New York: Addison-Wesley.
Cherniak, Christopher (1986), *Minimal Rationality.* Cambridge: MIT Press.
Dalla Chiara, Maria Luisa (1992) "Possible Worlds, Counterfactuals, and Epistemic Operators", in Christina Bicchieri and Maria Luisa Dalla Chiara (eds.) *Knowledge, Belief, and Strategic Interaction.* New York: Cambridge University Press, 155–166.
Doherty, Patrick (ed.) (1996), *Partiality, Modality and Nonmonotonicity.* Stanford: CSLI Publications.
Earman, John (1992), *Bayes or Bust?* Cambridge: MIT Press.
Edgington, Dorothy (1986), "Do Conditionals Have Truth-Conditions?", *Critica*: 18, 3–30.
Fetzer, James (1981), *Scientific Knowledge.* Dordrecht: D. Reidel.
——— and Donald Nute (1979), "Syntax, Semantics, and Ontology: A Probabilistic Causal Calculus", *Synthese* 40: 453–495.
Gärdenfors, Peter (1986), "Belief Revision and the Ramsey Test for Conditionals", *The Philosophical Review* 95: 81–93.
——— (1988), *Knowledge in Flux.* Cambridge: MIT Press.
——— (1992), "The Dynamics of Belief Systems: Foundations versus Coherence Theo-

also allows us to avoid the problem of comparing likelihoods of deterministic theories, all of which appear as if they must be 0 or 1 (the theory either entails the evidence or it does not). However, I presume that while it is true that $P(e|h_n) = 1$, when e is entailed by h_n where h_n is deterministic, it is not necessarily the case that $P(e|h_1 \,\&\, b) = 1$ and that $P(e|h_2 \,\&\, b) = 1$, where b is our background knowledge. Methodological, physical, and computational information in b may have a radical influence on our assessments of the likelihoods of those theories even if they are both formally deterministic.

24. Some aspects of this project have been worked out in Shaffer 2000.

ries", in Christina Bicchieri and Maria Luisa Dalla Chiara (eds.) *Knowledge, Belief, and Strategic Interaction.* New York: Cambridge University Press, 377–396.

Garey, Michael and David Johnson (1979), *Computers and Intractability: A Guide to the Theory of NP-completeness.* New York: W.H. Freeman.

Geroch, Robert and J. B. Hartle (1986), "Computability and Physical Theories", *Foundations of Physics* 16: 533–550.

Glymour, Clark (1981), *Theory and Evidence.* Chicago: University of Chicago Press.

Good, Irving J. (1983), "46656 Varieties of Bayesians", in his *Good Thinking.* Minneapolis: University of Minnesota Press, 20–21.

Hacking, Ian (1983), *Representing and Intervening.* Cambridge: Cambridge University Press.

Hájek, Alan (1989), "Probabilities of Conditionals Revisited", *Journal of Philosophical Logic* 18: 423–428.

Hanson, Norwood R. (1965), "Newton's First Law: A Philosopher's Door into Natural Philosophy", in Robert Colodny (ed.), *Beyond the Edge of Certainty.* New Jersey: Prentice-Hall, 6–28.

Hansson, Sven O. (1995), "The Emperor's New Clothes: Some Recurring Problems in the Formal Analysis of Counterfactuals", in G. Crocco, L. Farinas Del Cerro, and A. Herzig (eds.), *Conditionals: From Philosophy to Computer Science.* Oxford: Clarendon Press, 13–31.

Howson, Colin (1995), "Theories of Probability", *British Journal for the Philosophy of Science* 46: 1–32.

——— and Peter Urbach (1993), *Scientific Reasoning: The Bayesian Approach*, 2nd ed. Chicago: Open Court.

Hughes, R. I. G. (1990), "The Bohr Atom, Models and Realism", *Philosophical Topics* 18: 71–84.

Koyré, Alexander (1968), *Metaphysics and Measurement.* Cambridge: Harvard University Press.

Kyburg, Henry (1978), "Subjective Probability: Criticisms, Reflections, and Problems", *Journal of Philosophical Logic* 7: 157–180.

Langholm, Tore (1996), "How Different is Partial Logic?", in Patrick Doherty (ed.), *Partiality, Modality, and Nonmonotonicity.* Stanford: CSLI Publications, 3–34.

Laymon, Ronald (1989), "Cartwright and the Lying Laws of Physics", *Journal of Philosophy* 86: 353–372.

Leff, Harvey S. and Andrew Rex (eds.) (1990), *Maxwell's Demon: Entropy, Information, and Computing.* Princeton: Princeton University Press.

Levi, Isaac (1996), *For the Sake of the Argument: Ramsey Test Conditionals, Inductive Inference, and Nonmonotonic Reasoning.* Cambridge: Cambridge University Press.

Lewis, David (1976), "Probabilities of Conditionals and Conditional Probabilities", *Philosophical Review* 85: 297–315.

Lewis, David (1986), "Probabilities of Conditionals and Conditional Probabilities II", *Philosophical Review* 95: 581–589.

Lifschitz, Vladimir (1994), "Circumscription", in Dov Gabbay, C. Hogger, and J. Robinson (eds.), *Nonmonotonic and Uncertain Reasoning.* Vol. 3 of *Handbook of Logic in Artificial Intelligence and Logic Programming.* Oxford: Oxford University Press, 297–352.

Lindström, Sten and Wlodzimierz Rabinowicz (1989), "On Probabilistic Representation of Non-probabilistic Belief Revision", *Journal of Philosophical Logic* 18: 69–101.

——— (1990), "Epistemic Entrenchment with Incomparabilities and Rational Belief Revision", in A. Furhmann and M. Morreau (eds.), *The Logic of Theory Change.* Berlin: Springer-Verlag, 93–126.

——— (1992), "Belief Revision, Epistemic Conditionals, and the Ramsey Test", *Synthese* 91: 195–237.

——— (1995), "The Ramsey Test Revisited", in G. Crocco, L. Farinas Del Cerro, and A. Herzig (eds.), *Conditionals: From Philosophy to Computer Science.* Oxford: Clarendon Press, 147–191.

Makinson, David (1994), "General Patterns in Nonmonotonic Reasoning", in Dov Gabbay, C. Hogger, and J. Robinson (eds.), *Nonmonotonic and Uncertain Reasoning.* Vol. 3 of *Handbook of Logic in Artificial Intelligence and Logic Programming.* Oxford: Oxford University Press, 35–110.

McGee, Vann (1989), "Conditional Probabilities and Compounds of Conditionals", *The Philosophical Review* 97: 485–541.
McMullin, Ernan (1985), "Galilean Idealization", *Studies in the History and Philosophy of Science* 16: 247–273.
Milne, Peter (1997), "Bruno de Finetti and the Logic of Conditional Events", *British Journal For the Philosophy of Science* 48: 195–232.
Nowak, Leszak (1980), *The Structure of Idealization.* Dordrecht: D. Reidel.
Pitowsky, Itamar (1990), "The Physical Church Thesis and Physical Computational Complexity", *Iyyun* 39: 81–99.
Pitt, Joseph (1992), *Galileo, Human Knowledge, and the Book of Nature*. Dordrecht: D. Reidel.
Poole, David (1994), "Default Logic", in Dov Gabbay, C. Hogger, and J. Robinson (eds.), *Nonmonotonic and Uncertain Reasoning*. Vol. 3 of *Handbook of Logic in Artificial Intelligence and Logic Programming.* Oxford: Oxford University Press, 189–215.
Pylyshyn, Zenon (ed.) (1986), *The Robot's Dilemma.* New Jersey: Norwood.
Railton, Peter (1981), "Probability, Explanation, and Information", *Synthese* 48: 233–256.
Ramsey, F. P. (1990), *Philosophical Papers.* Edited by David H. Mellor. Cambridge: Cambridge University Press.
Redhead, Michael (1980), "Models in Physics", *British Journal for the Philosophy of Science* 31: 154–163.
Rosenkrantz, Richard (1981), *Foundations and Applications of Inductive Probability*. Atascadero, CA: Ridgeview Press.
Salmon, Wesley (1990), "Rationality and Objectivity in Science or Tom Kuhn Meets Tom Bayes", in C. Wade Savage (ed.), *Minnesota Studies in the Philosophy of Science*, vol. 14. Minneapolis: University of Minnesota Press, 175–204.
Shaffer, Michael (2000), *Idealization and Empirical Testing.* Ph.D. dissertation. Miami, FL: University of Miami.
Shapere, Dudley (1974), *Galileo: A Philosophical Study*. Chicago: University of Chicago Press.
Shoham, Yoav (1988), *Reasoning about Change*. Cambridge: MIT Press.
Stalnaker, Robert (1970), "A Theory of Conditionals", in William L. Harper, Robert Stalnaker, and G. Pearce (eds.), *Ifs*. London: Blackwell, 41–55.
Stich, Stephen (1990), *The Fragmentation of Reason*. Cambridge: MIT Press.
Walker, Elbert A. (1994), "Stone Algebras, Conditional Events, and Three Valued Logic", *IEEE Transactions on Systems, Man, and Cybernetics* 24: 1699–1707.