# Resolving the Bayesian Problem of Idealization

Nicholaos Jones

The Ohio State University

jones.1736@osu.edu

## 1 Introduction

Michael Shaffer (2001) challenges Bayesian confirmation theorists to show how at least some idealized hypotheses have at least some degree of confirmation. He argues that, in order to accomplish this task, one must either develop a coherent proposal for how to assign prior probabilities to counterfactual conditionals or abandon Bayesianism. This paper develops a Bayesian reply to Shaffer's challenge that avoids the issue of how to assign prior probabilities to counterfactuals by treating idealized hypotheses as abstract descriptions. The reply allows Bayesians to assign non-zero degrees of confirmation to idealized hypotheses and to capture the intuition that less idealized hypotheses tend to be better confirmed than their more idealized counterparts.

## 2 The Bayesian Problem of Idealization

According to Bayesian confirmation theory (hereafter: Bayesianism), the posterior probability of an hypothesis H given evidence E, Pr(H | E), determines how well E confirms H. (Bayesians typically interpret the function Pr(-) as a subjective probability.) And one hypothesis H1 is better supported by evidence E than rival hypothesis H2 just if E confirms H1 more than H2. That said, there is no general agreement among Bayesians on how to measure the degree to which evidence supports an hypothesis. Some Bayesians favor a difference measure, according to which the degree that E confirms H is equal to Pr(H | E) – Pr(H). Others favor a normalized difference measure, according to which the degree that E confirms H is equal to Pr(H | E) – Pr(H | ~E). Some favor a ratio measure: Pr(H | E) / Pr(H). And some favor a likelihood measure: {Pr(H | E)*[1 – Pr(H)]} / {[1 – Pr(H | E)]*Pr(H)}.

Despite these differences, all Bayesian measures involve the quantity Pr(H | E). This is the fact that is relevant to what Shaffer calls the Bayesian problem of idealization. The problem, in a nutshell, is this: Bayesian confirmation theory seems to entail that the posterior probability of every idealized hypothesis is undefined; this entails that Bayesianism is unable to account for the fact that some idealized hypotheses can be confirmed at least to some degree; and this entails that Bayesianism is unable to make sense of the intuition that less idealized hypotheses tend to be better confirmed than their more idealized counterparts.[1]

These unpleasant consequences result from the assumption that every idealized hypothesis is a counterfactual conditional in which the antecedent is a set of idealizing conditions -- such as "Each particle's radius $r \rightarrow 0$"[2] -- and the consequent is a set of claims that are true under those conditions -- such as the ideal gas equation $PV = NkT$. (More on this assumption in the next section.) If idealized hypotheses have the form A > C (where '>' is the symbol for counterfactual conditionals), then the posterior probability of an idealized hypothesis relative to evidence E has the form Pr(A > C | E). According to Bayes' Theorem, the posterior probability

---

[1] Shaffer further argues that since most scientific hypotheses are idealized, Bayesianism entails that "few, if any, scientific theories have ever been confirmed to any extent whatsoever" (2001: 45). This corollary makes Shaffer's argument more interesting. But I ignore it as incidental to the prior issue of whether Bayesianism can accommodate the fact that at least some idealized hypotheses are confirmed to at least some degree.

[2] I intend use of the arrow '→' to indicate that a limit is being taken -- in this case, the limit is one in which the particle radius $r \rightarrow 0$.

$Pr(A > C | E)$ is equal to $Pr(E | A > C)*Pr(A > C) / Pr(E)$. The problem for Bayesians is that there is no extant, coherent suggestion for how to assign prior probabilities to counterfactuals -- that is, for how to assign values to $Pr(A > C)$.[3] Since the posterior probability $Pr(A > C | E)$ is defined only if the prior probability $Pr(A > C)$ is defined, it seems that the posterior probabilities of idealized hypotheses are undefined. And, at least for Bayesians, this entails that the degree to which any given idealized hypothesis is confirmed by evidence is undefined. Given that most scientific hypotheses are idealized in some way, Bayesianism seems to entail that most scientific hypotheses cannot be confirmed.

Bayesians thus confront an apparent trilemma: either develop a coherent proposal for how to assign prior probabilities to counterfactuals; or embrace the counterintuitive result that idealized hypotheses cannot be confirmed; or reject Bayesianism. There is also a fourth option, developed in the remainder of this paper: reject the assumption that idealized hypotheses are counterfactual conditionals.

## 3 Motivating the Appeal to Counterfactuals

According to Shaffer, "When we claim that a theory holds in some idealized model, or under some idealizing conditions, we are claiming that a theory is true only on the basis of one or more *counterfactual* simplifying assumptions or conditions" (41). He insists that "theories incorporating idealizing conditions ought to be construed as counterfactuals" (43), adding that "this thesis is not open to question". Among others, Frederick Suppe (1989) and Ilkka Niiniluoto (1986) also claim that idealized hypotheses are counterfactuals. None of these authors motivates this claim. Perhaps it is worth pausing to fill the lacuna. The motivation seems to come in two

[3] Shaffer considers and rejects three suggestions. I endorse those considerations.

stages: the first stage motivates treating idealized hypotheses as conditionals; the second stage motivates treating these conditionals as counterfactuals.

*Stage One*. Claims obtained through appeal to idealizations tend to be false of real systems (despite sometimes being "close enough" to the truth for various purposes). The equation for the simple pendulum is false of most real pendulums, the ideal gas equation is false of most real gases, etc. If idealized hypotheses are the claims obtained through appeal to idealizations rather than conditionals in which such claims are the consequents, then most idealized hypotheses have a null posterior probability, in virtue of being inconsistent with available evidence. This result violates the intuitions that less idealized hypotheses tend to be better confirmed than their more idealized counterparts and that some idealized hypotheses have non-zero posterior probabilities relative to available evidence. Treating idealized hypotheses as conditionals rather than the claims obtained through appeal to idealizations allows them to be consistent with available evidence despite their consequents being inconsistent with that evidence.

*Stage Two*. If idealized hypotheses are conditionals (in which the antecedent is a set of idealizing conditions), then it is better to treat such conditionals as counterfactual rather than material. Since idealizing conditions typically are taken to be false, idealized hypotheses would be trivially true in virtue of the falsity of their antecedents if such hypotheses were material conditionals.[4] This result is unsatisfactory insofar as intuitions suggest that such hypotheses might be false. Moreover, saying that all idealized hypotheses are vacuously true in virtue of having false antecedents entails that the posterior probability of every idealized hypothesis is

[4] Leszek Nowak (1980: 136) recommends that, faced with the potential of idealized hypotheses being trivially true, we should continue to treat idealized hypotheses as material conditionals and revise the usual definition of truth.

unity relative to any evidence whatsoever. For most measures of degree of confirmation, this means that differences in the degree to which competing idealized hypotheses are confirmed depends entirely upon the prior probabilities of those hypotheses -- a counterintuitive result insofar as one expects evidence to play at least some role in determining rankings of the degrees to which competing hypotheses are confirmed.

**4 Idealized Hypotheses as Abstract Descriptions**

Until Bayesians develop a coherent proposal for how to assign prior probabilities to counterfactuals, and unless Bayesians want to deny that at least some idealized hypotheses can be confirmed, they should reject the treatment of such hypotheses as counterfactual conditionals. So, for example, they should reject treating the ideal gas law as an hypothesis of the form "If such and such idealizing conditions were to obtain, then the ideal gas equation would be true". They should reject treating the law of motion for simple pendulums as an hypothesis of the form "If such and such idealizing conditions were to obtain, then the equation of motion for the simple pendulum would be true". And so on.

At the same time, a Bayesian treatment of idealized hypotheses should not run afoul of the considerations raised in the previous section. A satisfactory Bayesian treatment should not entail that most idealized hypotheses have a null posterior probability. Nor should it entail that they are vacuously true. (This latter desideratum seems to require not treating idealized hypotheses as conditionals, in which case the ideal gas law just is the ideal gas equation and the law of motion for simple pendulums just is the equation of motion for the simple pendulum.)

All of these constraints can be met by treating idealized hypotheses as abstract descriptions. This proposal avoids the problem of how to assign prior probabilities to counterfactual conditionals, because it rejects the assumption that idealized hypotheses are counterfactuals.

An abstract description of a system is a partial or incomplete description of the system, a description that ignores certain features of the system but need not be false of the system. Such descriptions are commonplace. If someone says that the number of coins in his pocket is odd without saying anything else, his description of his pocket's contents is abstract in virtue of leaving aside details about how many coins are in his pocket. And if someone says that the gas inside the tube is a noble gas without saying anything else, her description of the tube's contents is abstract in virtue of leaving aside details about which noble gas is in the tube. None of these abstract descriptions are false.

This way of thinking about ordinary abstract descriptions is applicable to idealized descriptions. Let $f(x,y) = 0$ be an equation that is not idealized in any way and that correctly characterizes a physical system S. Let $g(x)$ be the function obtained by taking the idealizing limit of $f(x,y)$ in which $y$ approaches zero:

$$\lim_{y \to 0} f(x, y) = g(x),$$

so that $g(x) = 0$ characterizes an idealized version of S. Then there are two salient ways to understand the relation between the equation $g(x) = 0$ and S.

The equation $g(x) = 0$ can be understood as purporting to stand in a correspondence relation to S (or whatever relation the equation $f(x,y) = 0$ bears to S). Since the limit in which $y$ approaches zero is an idealizing limit, however, $g(x) = 0$ fails to stand in such a relation: $g(x) = 0$ is false of S, because the idealization used to obtain $g(x)$ from $f(x,y)$ is false of S. For instance, if $f(x,y) = 0$ is the equation of motion for a damped simple pendulum, if $g(x) = 0$ is the equation of motion for an undamped simple pendulum, and if the limit "$y \to 0$" idealizes the damping on pendulums, then $g(x) = 0$ is false of undamped simple pendulums because such pendulums are subject to more than an arbitrarily small amount of damping. This way of understanding the relation between an idealized description and the physical system it purports to characterize results from treating idealizations as falsities.

It is not mandatory to understand $g(x) = 0$ as purporting to stand in a correspondence relation to S but failing to do so. The same mathematical equation can be understood as standing in a correspondence relation to an abstract version of S rather than to S itself. Let $S^A$ be this abstract version of S, so that $g(x) = 0$ correctly characterizes $S^A$. Then the "$y \rightarrow 0$" idealization functions to transform one description -- viz., $f(x,y) = 0$ -- into a (more) incomplete description -- viz., $g(x) = 0$; and the "$y \rightarrow 0$" idealization determines which details $g(x) = 0$ ignores about S and the respects in which $S^A$ is an abstract version of S. Provided that there is an appropriate relationship between $S^A$ and S itself, it is possible for $g(x) = 0$ to be true of S because, as an abstract description of S, $g(x) = 0$ need not be false of S. (More on this below.) This point generalizes: an idealized description need not be false if it is an abstract description, because it only purports to characterize real systems indirectly, based upon whether the abstract system it characterizes bears an appropriate relation to real systems.

The preceding discussion highlights an important point, namely, that if idealized hypotheses are abstract descriptions, then the details they leave aside are determined by the idealizations used to obtain them. For instance, if the ideal gas law is an abstract description of real gases, then (among other things) it leaves aside details about interparticle forces. And if the law of motion for simple pendulums is an abstract description of real pendulum motion, then (among other things) it leaves aside details about the damping and pivot friction on real pendulums as well as the extension of real pendulum bobs.

In order for an idealized description of a system to be incomplete but not necessarily false, the idealizations used to obtain the description must ignore details about the system without attributing to the system features it does not have. In such a case, the idealization is said to be an abstraction rather than a distortion. Ernan McMullin (1985) provides a philosophical precedent for this way of understanding idealizations, noting that idealization "may involve a distortion of the original [system, description, etc] or it can simply mean a leaving aside of some components in a complex in order to focus the better on the remaining ones" (p. 248). McMullin's distinction between distortions and abstractions is a distinction between falsities and omissions: distortions falsify, whereas abstractions omit and need not distort.

In understanding idealizations as abstractions rather than as distortions, it is helpful to identify something as an idealization based upon its functional role. Idealizations have two characteristic functions. First, every idealization replaces a description of a system with a description of an idealized version of that system. Second, every idealization simplifies: the replacement description is, in some sense, simpler than the original description. For instance, working with the replacement description might simplify mathematical analysis. This way of classifying idealizations only appeals to the functional role of certain syntax; it does not appeal to a semantic interpretation of such syntax. And this allows the possibility of interpreting syntax that satisfies the appropriate functions as not purporting to represent the way the world is (e.g., as being an "inference ticket" rather than a truth-valued statement), and thereby allows the possibility that idealizations need not be false.

Consider a familiar idealized system, interpreting its characterizing idealizations first as distortions and then as abstractions in order to illustrate the difference between these interpretations. A damped simple pendulum is a pendulum that, among other things, is only subject to forces due to gravity and the damping of its surrounding medium (e.g., air). The damping tends to make the pendulum stop its oscillations. The behavior of a damped simple pendulum is correctly described by the following equation:

$$d^2\theta/dt^2 - b(d\theta/dt) + (g/L)\sin(\theta) = 0,$$

where θ is the angular displacement of the pendulum (this is a function of time), L is the distance from the pivot of the pendulum to its bob, g is the strength of the gravitational force, and b is the strength of damping.

In the idealized $b \to 0$ limit, the equation for the damped simple pendulum reduces to the equation for the simple pendulum:

$$d^2\theta/dt^2 + (g/L)\sin(\theta) = 0.$$

The mathematical role of the $b \to 0$ idealization is to transform the equation for the damped simple pendulum into the equation for the simple pendulum. This equation is to be understood as characterizing a pendulum in the limit where the amount of damping $b \to 0$.[5] There are (at least) two ways to interpret what the equation for the simple pendulum characterizes, one for each way of interpreting the $b \to 0$ idealization.

First, one might interpret the $b \to 0$ limit as a distorting idealization. Under this interpretation, the idealization says that the amount of damping on the pendulum is arbitrarily close to zero. And the equation for the simple pendulum characterizes a pendulum subject to a vanishingly small amount of damping. Even if the equation for the damped simple pendulum is

[5] The equation itself does not contain a term for damping; so this way of understanding the equation cannot be "read off" the equation itself. Nonetheless, the equation is about something. And what the equation is about, in part, is a pendulum for which the amount of damping $b \to 0$. The equation is also about a pendulum in which the friction at the pivot $F_f \to 0$, among other things. But I ignore these further complications, to keep the discussion simple and because they are not salient to the purpose of the example.

true of some real pendulums, the idealized equation for the simple pendulum, so interpreted, is false of all real pendulums.

Second, one might interpret the $b \to 0$ limit as an abstracting idealization. Under this interpretation, the idealization says that the amount damping on the pendulum is to be ignored (rather than made to be arbitrarily small). The equation for the simple pendulum provides a partial characterization of the damped simple pendulum, a characterization that ignores the amount of damping on the pendulum. Whereas the distortion-interpretation of the $b \to 0$ limit incorrectly represents the amount of damping on the damped simple pendulum by (incorrectly) attributing an arbitrarily small amount of damping to the pendulum system, the abstraction-interpretation of the same limit fails to represent the amount of damping on the pendulum by ignoring this feature of pendulum systems.

This example supports a more concise characterization of abstraction. According to Anjan Chakravartty, an abstraction is "a process whereby only some of the potentially many relevant factors or parameters present in reality are built-in to a model concerned with a particular class of phenomena" (2001: 327). According to Margaret Morrison, an abstract description is one that "does not include all of the systems [sic] properties, leaving out features that the systems [sic] has in its concrete form" (1999: 38 fn1). If idealizations are abstractions, then an idealization replaces one description of a system with a description that fails to attribute to the system at least one feature that the system has, without thereby attributing to the system a feature it does not have. For instance, if the equation of motion for the undamped simple pendulum is an abstract description of real pendulums and the idealizations used to obtain the equation are abstractions, then (among other things) those idealizations merely leave aside details about the damping and pivot friction on real pendulums. Likewise, if the ideal gas law is an abstract description of real gases and the idealizations used to obtain the law are abstractions, then (among other things) those idealizations merely leave aside details about interparticle forces.

The features ignored by an abstracting idealization are amounts of various properties. For example, as an abstraction, the idealization of damping on a particular pendulum ignores the amount of damping on that pendulum. It does not attribute an incorrect amount of damping to the pendulum. Nor does it say that there is a non-zero amount of damping on the pendulum, since ignoring the amount of damping is consistent with the pendulum having a zero amount of damping (e.g., swinging in a vacuum). As an abstraction, the idealization of damping simply does not specify the amount of damping on the pendulum; and in particular it does not specify an incorrect amount of damping.

In many cases, ignoring the amount of some property of a system effectively amounts to ignoring the property itself. For example, sometimes ignoring the amount of mass of a particle amounts to ignoring that the particle has mass at all. But this is not always the case. Consider the idealizing limit in which the particle mass $m \rightarrow 0$ for each particle in a system and the system's particle number $N \rightarrow \infty$ while the system's total mass $M = mN$ remains finite and non-zero. As an abstraction, this idealization ignores the amount of mass for every particle in the system but does not ignore the amount of mass of the system itself. Since the system having some mass entails that at least some of the system's components have mass, this idealization does not ignore the fact that some particles of the system have the property of mass.

Understanding idealizations as abstractions and idealized descriptions as incomplete has repercussions for the correctness conditions of idealized descriptions, because there is a sense in which partial descriptions can be "true" despite being incomplete. A reasonable assumption is that whether a partial description is true of a system depends upon whether what is ignored is "relevant" to the system.[6] (I take for granted some intuitive sense of "relevance".) For instance, since the equation for the simple pendulum ignores some features of the pendulum and sometimes partial descriptions can be "true", a reasonable assumption is that whether the equation for the simple pendulum is true depends upon whether the ignored features are "relevant" to the pendulum. Relevance is phenomenon-relative: a feature might be relevant with respect to one phenomenon of a system but not with respect to some other phenomenon of the same system, because not every phenomenon of a system always depends upon every feature of the system.

Since the correctness of a partial description depends upon the relevance of what is ignored, and since relevance is phenomenon-relative, it follows that an abstract equation is true of a system with respect to a given phenomenon of the system just in case what is idealized is not relevant to that phenomenon in that system. So, for instance, the equation for the simple pendulum is true of a real pendulum with respect to the rough, qualitative proportionality between the pendulum's period and its length just in case what is ignored in idealizing the pendulum (such as the amount of damping on it) is irrelevant to that proportionality. Hence, if idealizations are abstractions rather than distortions, the equation for the simple pendulum can be true of a real pendulum with respect to the rough proportionality between the pendulum's period and its length, even if there is not an arbitrarily small amount of damping on the pendulum. This

[6] There are alternative suggestions in the literature. For instance, Nancy Cartwright (1989; 1999) claims that equations that ignore some details of real systems describe capacities of abstract systems, and that when a real (concrete) system has the same capacities as an abstract system, the abstract equation describing that abstract system is also true of the real system. The suggestion here is more ontologically parsimonious than Cartwright's, since it does not postulate the existence of capacities.

possibility, of idealized equations being true of the systems they characterize, is absent if idealizations are distortions.

## 5 Solving the Problem

So much for explaining what it means to claim that idealized hypotheses are incomplete descriptions and that idealizations are abstractions. The remainder of this paper shows how endorsing these claims allows Bayesians to solve their problem with idealizations.

If idealizations are abstractions, then idealized hypotheses are not conditionals in which the antecedent is a set of idealizing conditions. For the antecedent of a conditional must be a set of statements; but idealizations are not statements if they are abstractions. Instead, they are more like "inference tickets" that transform one description of a system into a (more idealized) description that ignores certain features of the system. So treating idealizations as abstractions allows Bayesians to avoid worries about how to assign prior probabilities to counterfactual conditionals.

Moreover, if idealized hypotheses are abstract descriptions, there is nothing (especially) mysterious about how to determine their posterior probabilities via Bayes' Theorem. Prior probabilities are to be assigned to idealized hypotheses in the same way that such probabilities are assigned to incomplete or partial descriptions. And the conditional probability of the evidence given an idealized hypothesis (abstract description) will depend upon whether the features ignored by the hypothesis are relevant to the evidence. For instance, if the evidence shows only that there is a rough, qualitative proportionality between a specific real pendulum's period and the distance between its pivot and center of mass, then the conditional probability of this evidence given the law of motion for the simple pendulum is unity, since that law entails such a proportionality. But if the evidence also includes data about the exact period of a particular real pendulum, the conditional probability of this evidence given the law of motion for the simple pendulum is probably null, since most likely the equation for the simple pendulum predicts an incorrect period: it is probably false with respect to the exact period of the real pendulum in virtue of ignoring features that are relevant to the exact period.

Bayesians can avoid assigning zero as the conditional probability of the evidence given an idealized hypotheses by calculating this probability relative to a subset of all available evidence, such as evidence for which the features ignored by the idealized hypothesis are irrelevant. This selective attention to the evidence seems to accord with scientific practice. For instance, there is good reason to think that the current best scientific theories -- general relativity and quantum field theory -- are idealized. General relativity ignores quantum effects with the idealization that Planck's constant $h \rightarrow 0$, so that the Compton wavelength $\lambda_C = h/mc \rightarrow 0$. Quantum field theory ignores gravitational effects with the idealization that Newton's gravitational constant $G \rightarrow 0$, so that the Schwarzschild radius $r_S = 2GM/c^2 \rightarrow 0$.[7] But the failures of general relativity to accommodate quantum effects and of quantum field theory to accommodate gravitational effects are not taken to disconfirm those theories. Instead, the range of each theory is restricted to phenomena for which quantum or gravitational effects are irrelevant, respectively. And this restriction permits a restriction of the range of phenomena -- or sources of evidence -- that are eligible for confirming or disconfirming each theory. Similar restrictions occur with effective field theories, such as the Euler-Heisenberg theory for photon-photon scattering. This theory's range is restricted to phenomena in which the electron field is irrelevant to photon interactions (i.e., phenomena that occur at energy scales below the threshold

[7] The Compton wavelength is roughly "the distance scale at which quantum field theory becomes crucial for understanding the behavior of a particle of a given mass"; "the Schwarzschild radius is roughly the distance scale at which general relativity becomes crucial for understanding the behavior of an object of a given mass" (Baez 1999).

for electron production), so that the theory is not disconfirmed by phenomena in which the electron field is relevant to photon interactions.

Finally, if idealized hypotheses are abstract descriptions, it is possible to make sense of the intuition that less idealized hypotheses tend to be better confirmed than their more idealized counterparts. Consider the law of motion for the (undamped) simple pendulum and the law of motion for the damped pendulum. The latter is less idealized than the former, in virtue of taking into account the amount of damping on pendulums. So it is to be expected that set of evidence for which the features ignored by the law of motion for the damped pendulum are irrelevant is larger than the set of evidence for which the features ignored by the law of motion for the simple pendulum are irrelevant: the former set contains the latter plus evidence about phenomena in which damping is relevant. Relative to this larger set of evidence, Bayesians can expect the law of motion for the damped pendulum to be better supported than the law of motion for the simple pendulum, since the larger set of evidence is probably inconsistent with the predictions obtained from the law of motion for the simple pendulum.

**References**


Baez, John. (1999): Higher-Dimensional Algebra and Planck Scale Physics. In *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*, eds. Craig Callender and Nick Huggett. New York: Cambridge University Press: 177-198.

Cartwright, Nancy. (1989): *Nature's Capacities and their Measurement*. New York: Clarendon Press.

Cartwright, Nancy. (1999): *The Dappled World: A Study of the Boundaries of Science*. New York: Cambridge University Press.

Chakravartty, Anjan. (2001): The Semantic or Model-Theoretic View of Theories and Scientific Realism, *Synthese* 127(3): 325-345.

McMullin, Ernan. (1985): Galilean Idealization, *Studies in the History and Philosophy of Science* 16(3): 247-273.

Morrison, Margaret. (1999): Models as Autonomous Agents. In *Models as Mediators: Perspectives on Natural and Social Science*, eds. Mary Morgan and Margaret Morrison (New York: Cambridge University Press): 38-65.

Niiniluoto, Ilkka. (1986): Theories, Approximations, and Idealizations. In *Logic, Methodology and Philosophy of Science VII*, ed. Ruth Barcan Marcus (Amsterdam: North Holland): 255-289.

Nowak, Leszek. (1980): *The Structure of Idealization: Towards a Systematic Interpretation of the Marxian Idea of Science*. Dordrecht: Reidel.

Shaffer, Michael. (2001): Bayesian Confirmation of Theories That Incorporate Idealizations, *Philosophy of Science* 68, 36-52.

Suppe, Frederick. (1989): *The Semantic Conception of Theories and Scientific Realism*. Chicago: University of Illinois Press.