

AUTHOR'S POST-PRINT

Corrected for a typographical error as per the *Erratum*. Published version of the paper appears in *Philosophical Studies* (2009) 142 (3): 403-426, <<http://www.springerlink.com/content/901739151261uj4q/>>.

**Naturalism, Fallibilism, and the A Priori**

**Abstract.** This paper argues that a priori justification is, in principle, compatible with naturalism— if the a priori is understood in a way that is free of the inessential properties that, historically, have been associated with the concept. I argue that empirical indefeasibility is essential to the primary notion of the a priori; however, the indefeasibility requirement should be interpreted in such a way that we can be fallibilist about apriori-justified claims. This fallibilist notion of the a priori accords with the naturalist's commitment to scientific methodology in that it allows for apriori-justified claims to be sensitive to further conceptual developments and the expansion of evidence. The fallibilist apriorist allows that an a priori claim is revisable in only a purely epistemic sense. This modal claim is weaker than what is required for a revisability thesis to establish empiricism, so fallibilist apriorism represents a distinct position.

**Keywords:** Naturalism · Fallibilism · A priori · Epistemic possibility

**1. Introduction**

Naturalism is often thought to entail empiricism. Indeed, the two positions have sometimes been taken to be equivalent.<sup>1</sup> But they are conceptually distinct, and I will argue that in absence of further assumptions, the one does not imply the other. There are different naturalistic theses that are applicable to different domains, but in the context of epistemology, both the claim that the natural world is what exists (metaphysical naturalism) and the commitment to scientific method as the only legitimate means of attaining knowledge of the natural world (methodological naturalism) are central to a naturalistic orientation.<sup>2</sup> Although the ontological commitments of naturalism may be taken to be in conflict with a priori justification because of the latter's association with a realm of abstract eternal objects, it is methodological naturalism that poses the most direct

---

<sup>1</sup> See, for example, Devitt (1996), Maddy (2000), and Quine (1953).

<sup>2</sup> Various different types of naturalism are discussed in Audi (2000), Goldman (1999), and Stroud (1996). Different versions of epistemic naturalism are discussed in Goldman (1994) and Maffie (1990). See Kitcher (1992) for a formulation and critique of the traditional naturalist position in epistemology. See also De Caro and MacArthur, eds., (2004) for a collection of recent critical essays on contemporary naturalism.

challenge to the a priori. As Quine popularized the view, naturalism is “the abandonment of the goal of a first philosophy” and “the recognition that it is within science itself, and not in some prior philosophy, that reality is to be identified and described.”<sup>3</sup>

Empiricism might appear to follow from these methodological commitments in a few quick steps: If philosophy does not play a foundational role in inquiry by establishing a priori standards according to which science is to be evaluated, and if science – construed broadly as responsible empirical methods of belief formation and revision – is the only legitimate means by which we attain substantive knowledge, then it would appear that no statement could have (non-trivial) a priori justification conditions.<sup>4</sup> Scientific method is characterized *inter alia* by the ongoing responsiveness of theory to evidence, and recognizing some class of claims as immune to empirical evidence – should such evidence ever arise – would be in conflict with the naturalist’s commitment to scientific method.<sup>5</sup>

In this paper I will argue that a priori justification is, in principle, compatible with both metaphysical and methodological naturalism – if the a priori is understood in a way that is free of the inessential properties that, historically, have been associated with the concept. In Sect. 2, I identify the ways in which naturalism either is or might appear to be in conflict with a priori justification as it has traditionally been understood. I then present a naturalistic conception of the a priori in Sect. 3. I argue that a priori claims are empirically indefeasible; however, the indefeasibility claim should be interpreted in such a way that we can be fallibilist about apriori-justified claims. This fallibilist notion of the a priori accords with the naturalist’s commitment to scientific methodology in that it allows for apriori-justified claims to be sensitive to further conceptual developments and the expansion of evidence. Fallibilism and fallibilist apriorism

---

<sup>3</sup> Quine (1981, pp. 72 and 21, respectively).

<sup>4</sup> By a “non-trivial” form of a priori justification, I mean an a priori justification that cannot be fully explained in terms of analyticity or by stipulative definition.

<sup>5</sup> The rejection of a “first philosophy” is often understood as the denial that philosophy is an autonomous discipline, or equivalently in this context, that philosophy is continuous with science. Empiricism follows from the continuity thesis if it is combined with a strong version of confirmation holism: the seamlessness of the web of belief ensures that all claims are justified in the same way, namely via the coherence of the system as a whole with experience. (Scientific methods may presuppose a priori elements, so empiricism doesn’t follow from the continuity thesis alone.)

The denial that philosophy is an autonomous discipline is a separate and distinct thesis from the denial that it plays a foundational role – in the sense of providing infallible or rationally-irrevisable standards – in inquiry. For further discussion of this issue, see Siegel (1995). If the commitment to scientific method is taken to be sufficient for methodological naturalism, a naturalist might either reject the foundations thesis while upholding the autonomy of philosophy or reject strong confirmation holism (or both). The methodological naturalist would then not be rejecting a priori justification outright, but only claims of its incorrigibility or infallibility.

are further discussed in Sect. 4. Clarifying the fallibilist thesis with respect to a priori justification is a primary goal of this paper.

In Sect. 5, I identify some ways in which a suitably-naturalistic notion of the a priori either might be or has been employed to give naturalistic grounds for a priori belief. I do not offer a positive account a priori justification; my remarks in this section are intended merely to indicate how such an account (or accounts) might proceed. I close by considering some potential objections to a naturalistic understanding of the a priori in Sect. 6. Of particular concern is the worry that fallibilist apriorism collapses into a version of Quinean empiricism. I argue that if Quine is to be understood as offering an empirical alternative to apriorism, his revisability claim must be interpreted in a way that is stronger than mere fallibilism. Fallibilist apriorism is thus distinct position. If one adopts it, one need not reject the very idea of the a priori in order to be a thoroughgoing naturalist.

## 2. The Alleged Incompatibility of Naturalism and the A Priori

If naturalism encompasses the commitments of both metaphysical and methodological naturalism, then the a priori, as it has been traditionally understood, might be thought to be in conflict with naturalism for the following reasons:

- I. The subject matter of the a priori knowledge is Platonic.<sup>6</sup>
- II. Rational intuition, the alleged means by which we have knowledge of Platonic truths, is mysterious and occult.
- III. Non-trivial a priori claims are necessarily true, where the relevant notion of necessity is metaphysical.
- IV. Apriori-justified claims are rationally irrevocable, but scientific claims are, in principle, subject to revision in light of empirical evidence.

A number of contemporary philosophers have argued that the a priori should be understood along more modest lines whereby the properties identified in I - IV need not be included in our basic conception of the a priori.<sup>7</sup> Perhaps the most interesting and controversial aspect of these contemporary accounts is their treatment of a family of related notions: certainty, infallibility, and rational irrevocability. Naturalistically-minded philosophers disagree about whether there is any stable and philosophically-interesting notion of the a priori that can be isolated from these latter properties.

The term 'certain' is used in epistemological contexts in a number of different ways, and there are several different conceptions of certainty that might be invoked in characterizing a priori claims as certain. One widespread use of

---

<sup>6</sup> Naturalists who, following Quine and Putnam, take us to be committed to the existence of numbers (understood as sets) by virtue of the indispensability of mathematics to science will not object to the admission of abstracta *per se*, but only an endorsement of them absent a theoretical mandate.

<sup>7</sup> See, for example, Burge (1993; 1998), Field (1998; 2000), Goldman (1999) and Rey (1998).

‘certain’ that is not epistemic is the expression of the psychological notion of confidence: In the sentence ‘I am certain that he will finish the marathon race’, ‘I am certain that’ is roughly translatable as ‘I am highly confident that’. This psychological notion of confidence is not what is at issue in debates about whether a priori claims are known with certainty. Another conception of certainty that can easily be set aside is that of highest degree of warrant or credibility.<sup>8</sup> The a priori need not be associated with certainty in this sense, and moreover, it is implausible to think that a priori claims have a higher degree of warrant than claims about immediate sensory experience such as ‘I am now being appeared to redly’ or ‘I am now having a pain sensation’.

An important notion of certainty is that of absolute certainty or absolute infallibility. A propositional claim or belief is known with absolute certainty just in case it is not (genuinely) possible for the claim or belief to be mistaken. If it is not possible for the claim or belief that P to be mistaken, then it is not possible that P is false. Knowledge with absolute certainty thus entails the truth of its target claim or the propositional content of its target belief. Examples of claims that have been thought to be known with absolute certainty include *cogito*-type statements and statements, like those mentioned above, about immediate sensory experience that are relativized to an individual at a particular time. If a propositional claim, statement, or belief is thought to have the property of being known with absolute certainty, then presumably it has this property in virtue of a particular feature that it possesses.<sup>9</sup> Although one might want to make the case that an a priori warrant is a feature in virtue of which something is known with absolute certainty, knowing with absolute certainty is not part of the basic conception of the a priori.

There is, however, another notion of infallibility – more closely aligned with rational irrevisability – that might appear to be implicated in the basic a priori/a posteriori distinction. This notion pertains to the grounds or evidence that we have for a claim. If one is fallibilist about knowledge claims, one allows that (under contextually-relevant conditions) knowledge claims are correctable on conceptual grounds or upon the expansion of evidence. If P is thought to be known infallibly in this sense, we think we know that P will never come to be legitimately undermined, but we aren’t thereby taking P to be a kind of claim for which mistake is impossible. It is the notion of infallibility with respect to undermining evidence that is central to the question of whether a priori justification is compatible with naturalism: if a priori claims are understood to be empirically indefeasible, then it would appear that we are committed to claiming that we know that no empirical evidence will ever come to count against them.

---

<sup>8</sup> This notion of certainty is employed by Bertrand Russell in *Human Knowledge* (1948).

<sup>9</sup> A definition of absolute certainty would take something like the following form:

There is some associated class of beliefs  $Z(P)$  of which P is a member,  $\exists CP \leftrightarrow \neg \diamond(\text{Some member of } Z(P) \text{ is false})$ .

A criterion for membership in a relevant class Z would then need to be given to identify a class of absolutely certain beliefs. See Christopher Hookway’s recent discussion of fallibilism and the aim of inquiry (Hookway 2007) for discussion of a notion of absolute certainty that is developed along these general lines. See Firth (1967) for a taxonomy and discussion of epistemological uses of ‘certain’, including absolute certainty.

This would be inconsistent with methodological naturalism because it is decided in advance what the future evidence could be.

I take infallibility with respect to the correctness of a priori claims to be the central issue in establishing the compatibility of a priori justification with naturalism, and I will be articulating and defending a fallibilist notion of the a priori. I will argue that fallibilism about a priori justification can be understood as a special case of fallibilism about knowledge claims. I will not take up the question whether the susceptibility of our knowledge claims to modification in light of future conceptual advancements motivates either a cautionary stance toward belief or scientific antirealism; instead, I will restrict my discussion to showing how a priori claims can be coherently understood as fallible.

Given the overthrow of Euclid's parallel postulate and the discovery of the set-theoretic paradoxes, an apriorist has reason to be fallibilist about a priori claims independently of considerations of naturalism. A sensible apriorist should acknowledge that an a priori claim might be vulnerable to further conceptual developments – unless there are special reasons for thinking otherwise.

One can be fallibilist about an apriori-justified claim in two ways: one might allow that we could be mistaken either in thinking that a claim is true or in thinking that its justification conditions are a priori. In other words, one can be fallibilist about both *claims* that are said to be a priori warranted and the *a priori warrants* for the claims. If a proposition held to be a priori true were to come to be defeated on empirical grounds, both the proposition and our a priori warrant for it would be simultaneously overturned. However, a proposition might also come to be revised on *a priori* grounds. The need to revise naive set theory was recognized on purely conceptual grounds, and non-Euclidean geometries were developed prior to the discovery that space was non-Euclidean, so arguably the grounds for revision of an a priori claim could be purely conceptual. If an a priori claim were to be revised on purely conceptual grounds, the revision would not undermine its a priori status.

Likewise, we might come to discover that a proposition held to be true a priori is, in fact, empirically supported. In this imagined case, only our a priori justification, not the proposition itself, would be defeated. In the 19<sup>th</sup> Century, Carl Frederick Gauss doubted the a priori status of Euclid's parallel postulate, understood as a claim about physical space; however, he thought it to be empirically corroborated.<sup>10</sup> Had the parallel postulate been upheld on empirical grounds, only the warrant for the postulate, not the postulate itself, would have been undermined.

One caveat should be mentioned: although I will be articulating and defending a fallibilist notion of the a priori, it is important to recognize that a naturalistically-minded apriorist need not be committed to the fallibility of all a priori claims. There may be special reasons for thinking that a given a priori claim is known infallibly; however, the property of being known infallibly does not accrue to a claim solely by virtue of its a priori status.

I will argue that a suitably-modest and fallibilist notion of the a priori is, in principle, compatible with metaphysical and methodological naturalism.

---

<sup>10</sup> Gauss' skepticism regarding the a priori status of the parallel postulate and his conception of the possibility of alternative geometries is documented in his correspondence with Friedrich Wilhelm Bessel and H.C. Schumacher. The relevant portions of the correspondence are reprinted in Ewald (1996).

Although I won't argue for it here, it should also be clear from the discussion to follow that this notion of the a priori would likewise be compatible with naturalistic epistemologies.

### 3. A Modest A Priori

#### 3.1. The A Priori/A Posteriori Distinction

The a priori/a posteriori distinction is at root an epistemological distinction that contrasts two different types of justification. A priori justification may be characterized as follows:

**apj:** A priori justification is justification that does not rely upon the particulars of sensory experience for its justificational force.

The locution 'justificational force' is due to Tyler Burge.<sup>11</sup> It is intended to express the way in which a priori justification is typically understood to depend on experience in a general way, whereby experience may be needed to acquire the concepts utilized in an a priori claim, but the particulars of experience cannot constitute the evidence upon which the force of an a priori justification is based.

Although the a priori/a posteriori distinction is typically understood to contrast two types of justification, the distinction may be extended to contrast weaker forms of positive epistemic appraisal such as entitlement, reasonableness, or epistemically-blameless acceptance. The a priori/a posteriori distinction can arise for these weaker forms of positive epistemic appraisal as well as for justification. For example, it is often observed that no non-circular justification of very basic deductive and inductive inference rules can be given. If this is right, there is a *prima facie* problem for claims to their apriority: a principle cannot be apriori-justified if it cannot be justified at all. Some philosophers have argued that although we may not have a justification for our basic inference rules, we are nevertheless entitled to employ them – or at least we are epistemically blameless in doing so. If our acceptance of these rules is epistemically responsible independently of empirical experience (and the rules are empirically indefeasible), the basis for our acceptance of them should count as a priori even if it is not what we might regard as a full-fledged justification.

The a priori is often defined "in the negative" as I have done for several reasons. For one thing, there may be multiple forms of a priori justification. Alleged apriori-justified claims include not only the traditional subject matters of logic and mathematics, but also basic belief-forming methodologies, conceptual truths, second-order beliefs about the contents of occurrent first-order thoughts, Descartes' *cogito*, and the testimony of others in absence of reasons against it. It is not clear that any single positive characterization of a priori justification would encompass these diverse claims. Second, if the notion of the a priori is extended to weaker forms of positive epistemic appraisal such as entitlement or default reasonableness, the traditional notion that an a priori justification is "grounded in reason alone" may be inappropriately internalistic. Moreover, in those cases where we do take ourselves to be in possession of a reflective priori warrant, it is

---

<sup>11</sup> Burge (1993; 1998).

difficult to give a more perspicuous characterization of what it is for a warrant to be based on reason alone.

On the other hand, without further clarification of the role of sensory experience, a form of warrant such as memory that lacks an easily-identifiable sensory component might count as a priori according to the definition when intuitively it should not. While it is generally agreed that a more illuminating account of the a priori is needed, the prospects for providing such an account may depend upon our acquiring a more perspicuous understanding of the cognitive and physiological processes involved in experience.<sup>12</sup>

Regardless of the difficulties in drawing the a priori/a posteriori distinction, the empiricist cannot rest content with the charge that the distinction is unclear if only because she needs a notion of the a priori in order to deny that there is substantive a priori justification. And to the extent that we understand what an empirical justification is, we can contrast it with a non-empirical justification or a justification that is constituted by the operations of reason (or subdoxastic reasoning processes) upon previously-acquired concepts. So while I acknowledge that the notion of the a priori is in need of further clarification, I take the a priori/a posteriori distinction to be clear enough to support further refinements to the concept of the a priori and to sustain serious debate about whether any types of claims or principles are correctly classified as a priori.

### 3.2 A Priori Claims

Philosophers have been concerned with both *doxastic* or *personal* justification and with what might be called *propositional* or *impersonal* justification. Doxastic or personal justification specifies the conditions under which a particular agent S knows or is justified in believing that P. S may be justified in believing that P at a particular time and in a particular set of circumstances. Doxastic justification is thus individual-relative. By contrast, propositional or impersonal justification is not individual-relative: to say that a proposition is justified is to say that it is justified independently of whether any *particular* individual believes it. Justification at the impersonal level is predicated of propositions, rules, or subject matters.<sup>13</sup> An account of justification at the impersonal level will specify the evidence or reasons that can be given in support of a claim that is considered justified. In keeping with a naturalistic view of epistemology, it is important to understand the notion of impersonal justification as having its genesis in human thought: justification is not a property that some propositions possess independently of our judging them so.<sup>14</sup> Thus for a naturalistically-minded epistemologist, doxastic justification is the more basic notion. I will not be offering an account of how we move from doxastic justification to justification at

---

<sup>12</sup> A number of philosophers have emphasized the need for a more empirically-sophisticated account of experience in order to give a negative characterization of the a priori. See Casullo (2003, Ch.6); Goldman (1999); and Rey (1993; 2005).

<sup>13</sup> In what follows, I am not assuming that propositions have a mode of existence that is independent from the expressions we use to introduce them.

<sup>14</sup> Cf. Chisholm (1977). Chisholm seems to suggest that justification is an autonomous property that some propositions possess independently of our judgment of them.

the impersonal level, but presumably in any given case, it would involve the evaluation of a claim by some relevant subset of the members of an epistemic community.

In the discussion to follow, I will be talking about justification at the impersonal level. My primary reason for doing so is to evaluate a priori *claims* for their consistency with naturalism. When we say that a particular subject matter – for example, logic or mathematics – is a priori, we take ourselves to be fully aware of the subject matter’s alleged a priori status. Thus, I will be trying to establish the compatibility of *reflective a priori warrant* or *reflective a priori knowledge* with naturalism. Possessing reflective warrant at the propositional level is compatible with rejecting the “KK principle” (which says if S knows that P, then S knows that she knows that P) at the doxastic level. There may be conditions under which S knows that P, is justified in believing that P, or is entitled to infer in accordance with rule R, but her positive epistemic position vis-à-vis P or R is unknown to her. Apriorists and empiricists alike are concerned with the debate about which propositions or rules, if any, can be believed a priori, and in particular, which have a priori justification conditions.<sup>15</sup> While it would be of considerable interest to have an account of the conditions under which a cognitive agent might be justified in believing P or entitled to employ rule R when she is unaware of her right to do so, such an account would fall short of fully addressing what I take to be at issue between the apriorist and the empiricist.

So by focusing on a priori justification at the impersonal or propositional level, I intend to bypass important questions relating to the internalist or externalist character of justification and knowledge. These questions, which figure so prominently in discussion of doxastic justification and knowledge, are largely moot when the topic is reflective justification or knowledge. Thus, I will be operating with an internalist conception of knowledge and justification, but in so doing, I do not want to beg important questions at the doxastic level. If doxastic justification – or, perhaps more plausibly, doxastic entitlement – can be properly understood as externalist, one might want to reject a priori justification at the impersonal level, but allow that we have it at the doxastic level.<sup>16</sup>

What makes a given subject matter empirical or a priori? We classify something as empirical or a priori based on what we take its ideal justification conditions to be (ideal in the sense of ignoring individual limitations of computational capacity). Presumably, if we claim to have a priori knowledge of something, the a priori justification in favor of our knowledge claim is sufficient to underwrite it; additional empirical support is not needed. Likewise, if a propositional claim is taken to be known a priori, we would not expect empirical evidence to count against it. If we could foresee how a claim could be

---

<sup>15</sup> As Kripke (1980, p. 35) and others have noted, the claim that a particular proposition can be known a priori should not be understood as the claim that it must be known a priori. A reasonable apriorist will want to allow that a person could have empirical evidence for a proposition that can be known a priori. For example, a child just learning to count might combine five objects with seven objects, then count them to acquire evidence that  $5 + 7 = 12$ . But once the relevant concepts have been acquired,  $5 + 7 = 12$  could be calculated without the aid of objects.

<sup>16</sup> Externalist accounts of how we could have a priori justification or knowledge at the doxastic level have been advanced by Louise Antony (2004), Alvin Goldman (1999), and Georges Rey (1998; 2005).

undermined by empirical evidence, then we would take its justification conditions to be empirical, not a priori.

In light of the foregoing considerations, I take the claim that a proposition or rule of inference is a priori to include the requirement that it is empirically indefeasible. However, the empirical indefeasibility requirement needs to be stated in a way that will be consistent with fallibilism. A satisfactory formulation for these purposes is the following:

**Empirical indefeasibility:** A proposition (or rule of inference) is empirically indefeasible just in case no possible combination of observations should count against it.<sup>17</sup>

The ‘should’ in the definition requires some comment. There may be possible observations that could count against an a priori proposition or rule if we were to allow for an a priori claim to be sensitive to empirical evidence in inappropriate ways. For example, perhaps we could imagine taking a set of observations that conflicted with the predictions of a test hypothesis to undermine our hypothetico-deductive method rather than giving us a reason to either reject the hypothesis or modify at least one of our background assumptions. But this would obviously be epistemically irresponsible. (This is not to claim that the elementary inference rules implicated in our H-D method such as modus ponens and modus tollens are unassailable,<sup>18</sup> but simply to point out that making an a priori inference rule or claim answerable to experience in ways that are patently unattractive ought not to count against its a priori status.)

On the other hand, there are cases for which there is (or has been) genuine disagreement about whether empirical observations constitute evidence against an allegedly a priori claim. The proposal to revise a version of the distributive laws (in one direction) in light of the paradoxes of quantum mechanics is – or at least may have been – a case in point. If the hypothesis that the distributive law doesn’t hold in the quantum domain were to be regarded as a live hypothesis, whether the observed paradoxes should be judged to undermine the distributive law would depend on theoretical considerations such as whether apparent violations of the distributive law are best explained by faulting the logic and whether adopting a different logic would solve the problem. A judgment about the evidential relation of observation to hypothesis in this case would presumably be resolved in the same manner as other theoretical scientific disputes.

As a first approximation then, the notion of a priori insofar as it might be used to classify a particular subject matter that I want to defend is as follows:

---

<sup>17</sup> This is Hartry Field’s formulation (Field 1998) of empirical indefeasibility in his discussion of the apriority of logic.

<sup>18</sup> Alleged counterexamples to modus ponens have been put forth by Adams (1975), McGee (1985), and Lycan (1993); and some philosophers have taken the sorites paradoxes to show that modus ponens is not a strictly valid form of inference. (Whether any particular challenge is an empirical challenge is a further question.) Evaluating these challenges would *inter alia* involve interpreting the role of idealization assumptions in the formalization of deductive rules and settling questions about the topic neutrality of logic.

**app:** A proposition is a priori iff (1) it can be justifiably believed without empirical evidence, and (2) it is empirically indefeasible, if so justified.

The definition of an a priori proposition can be extended to a system of rules as follows:

**apr:** A rule or system of rules is a priori iff (1) it can be justifiably employed without empirical evidence, and (2) it is empirically indefeasible, if so justified.

The extension of the a priori/a posteriori distinction to a system of rules is important to the question whether scientific methods presuppose a priori elements.

#### 4. Fallibilism

Although a number of contemporary philosophers have advocated a fallibilist notion of the a priori, most have not attempted to spell out what fallibilism amounts to in this context in any detail.<sup>19</sup> (An exception is Hartry Field (1998; 2000). I discuss a challenge that Field poses for interpreting the fallibilist thesis below.) A fallibilist about a priori justification thinks that an a priori claim is empirically indefeasible, but he allows that we don't know that future conceptual developments will never come to reveal empirical evidence against it.

Suppose we think the distributive law  $p \wedge (q \vee r) :: (p \wedge q) \vee (p \wedge r)$  holds true and is a priori, but we want to be fallibilist about a priori justification. If further advancements in physics that we cannot now anticipate eventually come to vindicate a version of quantum logic that does without the distributive law in one direction, we will have been mistaken in thinking this version of the distributive law is a priori true. Although we don't think there are any observations that we might make that should count against the distributive law, it may simply be a failure of our imagination to see how empirical evidence should count against it. The recognition that we might be mistaken in thinking that no empirical evidence should count against the law does not undermine its (alleged) a priori status: our fallibility is not evidence for or against the law; *a fortiori*, it is not evidence against the law's being a priori justified. The fallibilist merely allows that we don't know that an a priori claim will never come to be undermined by future conceptual developments. (A 'conceptual development' here should be understood broadly enough to include new discoveries made by more rigorous applications of existing methods of analysis. Russell's discovery of the set theoretic paradoxes was a conceptual development on this understanding of the term.)

The fallibilist interpretation of the empirical indefeasibility requirement is analogous to the way in which we might want to allow that knowledge claims, generally, may be vulnerable to the expansion of evidence. Unless we take our

---

<sup>19</sup> Albert Casullo (2003, Ch.3), distinguishes between two notions of fallibility with respect to a priori justification that are sometimes conflated: a notion of fallibility that allows for an apriori-justified belief to be false, and a notion of fallibility according to which a belief is defeasible. No entailment between a priori justification and truth is asserted on my modest notion of the a priori, so it is only Casullo's second sense of fallibility that needs to be considered.

fallibility as a reason to adopt a cautious stance toward our beliefs, allowing for the defeasibility of knowledge claims should not be taken to undermine the knowledge that we take ourselves to possess: what is claimed as known *is* known according to ordinary standards for knowledge, our current understanding, and all of the available evidence. The defeasibility condition allows that, for all we know, a knowledge claim might come to be overturned upon further conceptual developments and the ensuing expansion of evidence. Analogously, to allow that the apriority of a claim might be *empirically* defeasible in light of future conceptual developments and observations that we might make in light of them should not undermine its a priori status according to our existing body of knowledge.

#### 4.1 Pure Epistemic Possibility

A fallibilist interpretation of empirical indefeasibility might appear to have an air of contradiction about it: wouldn't being fallibilist about empirical indefeasibility undermine the claim that any proposition or rule is genuinely a priori justified? If it is possible that empirical evidence could come to count against an a priori claim, then it might appear that the claim is revisable in light of experience, and so its justification conditions are not a priori after all. As Field (1998) puts the point in his discussion of the apriority of logic:

if one believes it possible that logic will turn out to be empirically revisable, one surely must believe it possible that logic will be empirically revised; but that is to believe logic empirically revisable, and so not a priori (p. 4).

This argument assumes the principle  $\diamond\diamond \rightarrow \diamond$ , which Field notes is hard to fault in the current context. His diagnosis of why the argument fails is that it turns on a conflation of two kinds of possibility: the fallibilist about apriorism allows for the “purely epistemic” possibility that an a priori claim will come to be empirically undermined, but this purely epistemic possibility does not entail logical (or genuine) possibility – which explains why it does not obey the usual modal axioms.

How should this notion of pure epistemic possibility be understood? I think it would be a mistake to think of pure epistemic possibility as a separate and distinct kind of possibility. An epistemic possibility is an agent-relative notion that obtains in virtue of a body of knowledge or evidence available to a cognitive agent (or community of agents).<sup>20</sup> A pure epistemic possibility can be defined in terms of knowledge operators. I propose to define the basic notion of a pure epistemic possibility as follows: something is possible in a purely epistemic sense

---

<sup>20</sup> This general understanding of epistemic possibility is found in Peirce, who took epistemic possibility to be our most basic employment of the term. See Peirce's *Collected Papers*, Vol. 6, Section 367, 1901, and Vol. 5, Section 454, 1905. I am grateful to Robert Meyers for bringing the relevant passages in Peirce to my attention.

just in case it is not known not to be true. Symbolically,  $\diamond_{\text{pure epistemic}} = \neg K \neg$ . A pure epistemic possibility thus defined does not entail genuine possibility.<sup>21</sup>

To say that a proposition is known is shorthand for saying that some person(s) knows it. Statements about what is or isn't known within a community also invoke social standards of evidence and confirmation. When we assert, 'It is purely epistemically possible that P', this epistemic claim must be understood as a claim about what is not known in an epistemic community. (Again, I will set aside questions concerning how and under what conditions a proposition becomes established in an epistemic community.) For the purposes of discussing the status of scientific claims or general a priori claims, the epistemic community can be understood to be everyone (or at least everyone here on earth), although, of course, it need not be the case that each person in the community be in a position to assess a given claim. On the other hand, it would not count against S's assertion 'It is not known that not P' if a hypothetical community that was completely inaccessible to S did, in fact, possess knowledge that not P. The truth conditions for pure epistemic possibility ("PEP") are as follows:

**PEP:** S's assertion 'It is purely epistemically possible that P' made at  $t$  is true iff, at  $t$ ,  $\neg P$  is not known by a contextually-relevant subset of members of an epistemic community that is, in principle, accessible to S].

Pure epistemic possibility as defined above encompasses two established epistemic uses of 'possibility.'<sup>22</sup> The first is the expression of ignorance with respect to P for cases where we don't have adequate grounds for either accepting or rejecting P. Goldbach's conjecture is epistemically possible in this sense: we can't yet say whether it is true or false. The second notion, and the one with which I am ultimately concerned, is the epistemic possibility that characterizes fallibilism about knowledge claims: although we think we know P, we allow for the purely epistemic possibility that we are mistaken about P because we are imperfect cognitive agents.

Fallibilist apriorism is a claim about what is not known with certainty about a genuine possibility, viz. the possibility that an apriori-justified claim will come to be revised on empirical grounds. The notion of genuine possibility here need not be understood as "metaphysical possibility," which, as I indicated in Section

---

<sup>21</sup> Pure epistemic possibility as I am defining it corresponds to what Tamar Szabó-Gendler and John Hawthorne characterize as a *permissive* notion of epistemic possibility. By contrast, a *strict* account of epistemic possibility would entail metaphysical possibility. An example of a strict notion of epistemic possibility would be the following: P is epistemically possible for S just in case P is metaphysically compossible with all that S knows ( Szabó-Gendler and Hawthorne 2002, pp. 3-4).

<sup>22</sup> An additional epistemic use of possibility is the notion of evidential neutrality that Kripke introduced to express the sense in which a necessary a posteriori claim could have turned out to be false. 'Hesperus' and 'Phosphorus' name one and the same object, so it is not possible that Hesperus  $\neq$  Phosphorus; yet, given the evidence, it could have turned out that Hesperus and Phosphorus were two distinct objects. However, this notion of epistemic possibility does entail a genuine possibility, so it is not pure as defined above. See George Bealer (2002, Section 1.3) for a recent discussion of different epistemic notions of possibility. For further discussion of Kripke's notion of epistemic possibility, see Stephen Yablo (1993: 22-25; 2002).

III, might be objectionable to a naturalist. If the notion of metaphysical possibility is objectionable, an alternative would be to understand possibility as logical possibility plus a demand for consistency with a contextually-relevant condition such as a physical law, an accepted theory, or a socially-instituted practice. So, for example, it would be possible for meteor on a set path to collide with the earth iff it is physically possible. It is possible that Goldbach's conjecture is true iff its being true is consistent with mathematical theory. It is possible that a prisoner on death row will be granted a stay of execution on appeal iff the reasons for his request merit serious consideration under the prevailing laws. There may be disagreement over which conditions are in play with respect to the evaluation of a particular possibility claim, but a notion of possibility that is not merely formal need not be understood as metaphysical possibility.

In the discussion to follow, I operate on the assumption that there could be statements that are true but not knowable by us. However, to the extent that the relevant distinctions between epistemic and genuine possibilities could be drawn within a verificationist framework, one should be able to articulate the fallibilist thesis that I develop below within that framework. (I do not attempt to do this here.)

#### 4.2 Pure Epistemic Possibility vs. Epistemic Analyses of Possibility

Pure epistemic possibility as I am defining it should also be distinguished from the epistemic sense of possibility that has been advocated by some philosophers as the correct analysis of the possibility expressed in sentences of the form 'It is possible that P', where the embedded P is in the indicative mood.<sup>23</sup> These epistemic accounts of possibility attempt to state truth conditions for S's assertion 'It is possible that P' in terms of what S does not know and what is or is not *knowable*. Epistemic possibility in this sense entails genuine possibility. The challenge for a fully epistemic account of possibility is to give an account of possibility that is extensionally correct without relying upon non-epistemic notions.<sup>24</sup>

An epistemic account of possibility that entails genuine possibility cannot have the truth conditions for PEP given above because something can fail to be

---

<sup>23</sup> The epistemic analysis for the possibility expressed by sentences in the indicative mood is defended by DeRose (1991; 1998), Hacking (1967; 1975), G.E. Moore (1962), and Teller (1972). The authors differ in the details. Hacking also identifies a sense in which possibility is epistemic in subjunctive cases: It is logically and perhaps causally possible that I should have been blind by now, but it is not possible that I am blind now, though it is possible that I shall be blind tomorrow. 'It is possible that I shall be blind tomorrow' expresses my state of not knowing that I will not be blind tomorrow, but in order for me to have said something true, it must also be genuinely possible for me to be blind tomorrow.

<sup>24</sup> Hacking proposes (but later retracts) an analysis of epistemic possibility according to which a state of affairs is possible if it is not known not to obtain, and no practicable investigations would establish that it does not obtain. Other proposals, e.g. DeRose (1991), appeal to what *can* be known to solve this problem. I am skeptical, however, that an account can be given that, when pressured, doesn't invoke a truth-theoretic notion of possibility.

(genuinely) possible, despite the fact that no one knows it is false. Ian Hacking (1967) illustrates the problem with his *Salvage Ship Case*:

Imagine a salvage crew searching for a ship that sank a long time ago. The mate of the salvage ship works from an old log, makes some mistakes in his calculations, and concludes that the wreck may be in a certain bay. It is possible, he says, that the hulk is in these waters. No one knows anything to the contrary. But in fact, as it turns out later, it simply was not possible for the vessel to be in that bay; more careful examination of the log shows that the boat must have gone down at least thirty miles further south. The mate said something false when he said, “It is possible that we shall find the treasure here,” but the falsehood did not arise from what anyone actually knew at the time (p. 148).

While the mate’s statement “It is possible that we shall find the treasure here” was false, it would not have been false to say that it was purely epistemically possible for the mate that the treasure was in the bay. It was not known to the mate and others that the ship could not have been in the bay, yet it could not have been; it was purely epistemically possible – possible for all the mate knew – but not actually possible.

#### 4.3 Fallibilism and a Fallibilist A Priori

A fallibilist apriorist thinks an a priori claim is empirically indefeasible, but he allows that in a certain purely epistemic sense, it is possible that an a priori claim will turn out to be empirically defeasible after all. When the fallibilist asserts ‘It is (purely epistemically) possible that an a priori claim will turn out to be empirically defeasible’, he asserts the proposition that is generated by the application of the propositional operator  $\diamond_{\text{pure-epistemic}}$  to the proposition that  $P_{\text{apj}}$  will turn out to be empirically defeasible. Symbolically,  $\diamond_{\text{pure-epistemic}} [P_{\text{apj}}$  will turn out to be empirically defeasible]. Or equivalently,  $\diamond_{\text{pure-epistemic}} \diamond_{\text{genuine}} [P_{\text{apj}}$  will come to be empirically defeated]. In what follows, I will drop the subscript ‘genuine’ for genuine possibility.

If ‘ $\diamond_{\text{pure-epistemic}}$ ’ is analyzed as ‘ $\neg K \neg$ ’, which in the current context can be read as ‘it is not known not to be the case that’,<sup>25</sup> then it might appear that fallibilist apriorism could be represented as

$$(1) \quad \neg K \neg \diamond [P_{\text{apj}} \text{ will come to be empirically defeated}].$$

Because  $\neg K \neg \diamond$  does not imply  $\diamond$ , the actual empirical defeasibility of  $P_{\text{apj}}$  does not follow from the fallibilist thesis.

---

<sup>25</sup> Strictly speaking, the use of a pure epistemic possibility operator requires reference to one or more knowers. When S asserts ‘It is not known not to be the case that P’, S asserts that no contextually-relevant subset of members of his epistemic community, where the contextually-relevant subset may include S, knows that not P at the time of his assertion. Where ‘ $\Gamma$ ’ is an epistemic community for S, S asserts a statement of the form  $\neg(\exists x)[(x \in \Gamma) \ \& \ Kx(\neg p)]$ . For ease of exposition, I will be suppressing reference to both the asserters and the knowers, except when it is important to identify them.

However, (1) does not adequately express fallibilist apriorism: in asserting (1), we assert that we don't know what we take ourselves to know, namely that  $P_{apj}$  is empirically infeasible, and so it is *not* the case that it is genuinely possible that  $P_{apj}$  will come to be empirically revised. In general, if pure epistemic possibility is understood without further qualification as 'it is not known not to be the case that', we don't distinguish fallibilism with respect to P from ignorance of P.

I think the purely epistemic notion of possibility can be used to characterize fallibilist apriorism, but this will need to be done in such a way that captures the particular circumstances in which the fallibilist allows that an a priori claim might come to be empirically revised. I will approach this task by first considering how fallibilism in general should be understood, then considering the particular case of fallibilist apriorism. In what follows, I will not be defining a knowledge operator within a formal system, nor will I be giving a semantics for 'know'; I will be giving an explanation of how the purely epistemic notion of possibility implicated in the fallibilist thesis should be understood.

Suppose we think we know that P, but we are fallibilist about knowledge claims; that is, we think it's a purely epistemic possibility that the claim that P is defective. As (1) above revealed, this fallibilist claim cannot be expressed simply as

$$(2) \quad \neg K\neg[\neg P] \text{ (or equivalently, } \neg KP)$$

because when P is thought to be known, we would simultaneously be asserting KP and  $\neg KP$ .

Statement (2) exhibits a familiar problem for stating the fallibilist thesis: it is difficult to give a statement of the position that doesn't lead directly to skepticism.<sup>26</sup> Of course in cases where P is thought to be known, the fallibilist wants merely to claim that we can't dismiss the *possibility* of P's coming to be undermined. So it might be thought that the fallibilist thesis could be stated as

$$(3) \quad \neg K\neg[\diamond\neg P] \text{ (or equivalently, } \neg K\square P)$$

But (3) sounds like we are uncertain of the modal status of P: If P is known to be a contingent claim, it would be both misleading and pointless to assert (3), which is equivalent to  $\neg K\square P$ . On the other hand, if P is a necessary statement, (3) asserts that we don't know that it is. So for cases where P is known to be a necessary claim, the fallibilist is once again in the position of asserting a contradiction, this time concerning the necessity of P. Moreover, (3) doesn't express fallibilism with respect to P itself.

The temptation to express fallibilism as a claim about what is not known about a genuine possibility as in (3) above is due to a failure to recognize that the possibility implicated in the fallibilist thesis is fully analyzed as a pure epistemic possibility. In an attempt to state the fallibilist thesis in a way that doesn't contradict the claim that P is known, the thought was that some further possibility – the genuine possibility that not P – is being expressed in the fallibilist claim. But the result was an assertion of ignorance with respect to the modal status of P.

---

<sup>26</sup> This difficulty is explored in Lehrer and Kim (1990).

The sense in which the fallibilist allows that P might fail to hold cannot be expressed as a straightforward claim about the modality of P.<sup>27</sup>

The key to a correct statement of fallibilism with respect to knowledge claims at the impersonal level lies in recognizing that it is a second-order claim about the susceptibility of our beliefs and knowledge claims to error in a specific context of evaluation.<sup>28</sup> The fallibilist recognizes that we are imperfect cognitive agents and that, therefore, our knowledge claims may be correctable. (The exception would be those claims, if any, that are known with absolute certainty because they belong to class of claims for which it is not genuinely possible that any member of the class be mistaken. In what follows, this class of claims should be understood to be excluded from the discussion.)

For knowledge claims of a general sort that are not relative to an individual at a particular time, individual computational errors and failures to consider readily-available alternatives are largely discounted by the vetting process required for the claim to become accepted by a relevant epistemic community. So the potential source of error for an established knowledge claim that the fallibilist wants to acknowledge is the limitation of our current understanding with respect to its evaluation. In so doing, the fallibilist is not expressing doubt about a particular claim, or class of claims; she is merely acknowledging the relevant way in which general knowledge claims are not immune to defeasance or modification. Fallibilism thus understood is a view about our knowledge claims that recognizes both our limitations as knowers and the dynamism of our corpus of beliefs.

The fallibilist wants to acknowledge that we don't know that P will never come to be legitimately undermined. To express the fallibilist thesis, instead of (2) or (3) above, we want something like

- (4)  $\neg K\neg[P \text{ will come to be legitimately undermined in the future}]$ .

As it stands, when P is thought to be known, the assertion of (4) retains an air of contradiction: If P is known, then P is true. But if P is true, then it follows that P will not come to be legitimately undermined. So it would appear that the fallibilist cannot both claim that P is known and that it is *not* known that P will not come to be legitimately undermined in the future (or, more colloquially, that we don't know that P will never come to be legitimately undermined).

However, different standards for knowledge are in play when the fallibilist asserts 'It is known that P' and 'We don't know that P will never come to be legitimately undermined'. If we knew that P would never come to be legitimately undermined, we would know P in a *context of assessment* that takes into consideration future conceptual advancements and observations that we might make in light of them.<sup>29</sup> This context of assessment specifies a higher standard for

---

<sup>27</sup> The mistake in construing fallibilism as a claim about modality is discussed by Susan Haack (1979) and Robert Meyers (1988), among others.

<sup>28</sup> That fallibilism is a second-order claim has been emphasized by Simon Evnine (2001) and Jonathan Adler (2002) in their discussions of fallibilist belief.

<sup>29</sup> The notion of a context of assessment has been introduced recently in connection with a form of contextualism, dubbed 'relativism', that takes 'knows' to be sensitive to a context of assessment, as opposed to the context of speaker use or the circumstances (e.g. time) of evaluation of a knowledge claim. On a relativist theory of knowledge-attributing

knowledge than what is in play for ordinary knowledge claims. Knowledge of P according to this higher standard would be infallible knowledge. When the fallibilist asserts 'We know (or don't know) that P will never come to be legitimately undermined', the higher standard is invoked via the assertion.

Let  $K^*[P]$  be knowledge in the context of assessment that gives the higher standards for infallible knowledge of P:

- (5)  $K^*[P]$  =<sub>df</sub> knowledge in a context of assessment for P that has been expanded to take into consideration future conceptual advancements, imaginable or otherwise, that we might make, and observations that we might make in light of those advancements.

A context of assessment is a set of circumstances and standards according to which P may be epistemically evaluated. The expanded context of assessment in (5) is a qualitative change to a context of assessment that raises the standards for knowledge beyond what is required for ordinary knowledge claims.

An ordinary context of assessment for a claim of a general and theoretical nature would be determined, for the most part, by the norms of evaluation of those epistemic agents who are in a position to evaluate it. When S asserts that P is known, S may be claiming to know P directly in accordance with the appropriate norms of evaluation for P, or S may be deferring to a group of experts in his community. In the latter case, S's claim to know P is based on testimony, and the relevant standards of assessment for P are set by the group of experts to whom S is deferring. For example, if S claims to know that  $2 + 3 = 5$ , S likely represents himself as being familiar with elementary arithmetic. However, if S doesn't know much about neurophysiology, but he has heard from reliable sources that visual processing takes place in the occipital lobe, S defers to the neuroscientists who are in a position to evaluate the truth or falsity of the claim.

In asserting fallibilism with respect to P, the fallibilist invokes a higher standard for knowledge,  $K^*[P]$ , that is independent of any particular scientific discipline or ordinary context of evaluation. However, the fallibilist does not claim knowledge of P when these higher standards are in place: If P were known according to the higher standards for  $K^*[P]$ , it would not be epistemically possible for P to come to be justifiably undermined upon further conceptual advancements because it would be known that P will never come to be so undermined.

Fallibilism with respect to P can be expressed as

- (6)  $\neg K^*[\neg P]$  (or simply  $\neg K^*[P]$ )

The fallibilist is not in the position of asserting a contradiction when she asserts that P is known, yet it is possible in the purely epistemic sense that P should come to be undermined by future conceptual developments because, in effect, she asserts  $KP$  but  $\neg K^*[P]$ .

---

sentences, the truth of a statement employing an epistemic modal is relative to a context of evaluation. Different versions of contextual relativism are defended by John MacFarlane (2005a; 2005b); Mark Richard (2004); Andy Egan, John Hawthorne, and Brian Weatherson (2005); and Andy Egan (2007).

With a general statement of fallibilism in hand, we now have the resources to state fallibilist apriorism. Fallibilism with respect to empirical indefeasibility can be distinguished from ignorance about it using the knowledge operator  $K^*$ :

(7)  $\neg K^* \neg [P_{\text{apj}} \text{ is empirically defeasible}]$ ,

or equivalently,

(7')  $\neg K^* \neg \diamond [P_{\text{apj}} \text{ will come to be empirically defeated}]$ .

The knowledge operator in (7) and (7') specifies a standard for knowledge according to which we would not claim to know that a claim that we take to be a priori is not empirically revisable. The context in which this standard is invoked is the same as the context in which a fallibilist acknowledges that we don't know that an established knowledge claim is defective – namely the context in which future conceptual advancements and observations that we might make in light of those advancements are taken into consideration. Fallibilist apriorism is just a particular instance of the more general case.

## 5. The Compatibility of a Modest A Priori with Naturalism

The modest notion of the a priori outlined above is immune to the most of the objections that I raised in Section 2 on behalf of the naturalist. There is no conflict with metaphysical naturalism (and hence attempts to reconcile epistemology with metaphysical naturalism) because no commitment is made either to Platonic entities as the subject matter of the a priori or to rational intuition as the source of a priori justification. The aspects of a priori justification that are most objectionable to the methodological naturalist are likewise largely eliminated on a fallibilist notion of the a priori. A priori justification need not be associated with any particular view about metaphysical necessity. Nor are the properties of absolute certainty, rational irrevisability, or infallibility attributed to a priori claims. A fallibilist apriorism allows for the defeasibility of an apriori-justified claim in the same way that knowledge claims, generally, may be understood to be defeasible. Thus objections I – IV in Section 2 do not apply to a suitably-modest notion of the a priori.

However, if rational intuition is rejected, one might wonder what other a priori account is on offer.<sup>30</sup> I will not attempt here to give a naturalistic account of a priori justification – either for belief or for rules of inference, but I will identify some recent approaches to providing the desired naturalistic grounds. My own view is that it is highly unlikely that there will be a single, unified naturalistic account of a priori. My reasons for thinking this are as follows: First, as I noted

---

<sup>30</sup> It should be pointed out that an apriorist who wanted to retain rational intuition as a source of a priori knowledge could give a naturalistic account of rational intuition itself. A naturalistic construal of rational insight or intuition might characterize it as a (noninferential) judgment of evidentiality with no attendant claim about access to an independent realm of truths. The term “rational insight” thus might be understood only as loose talk meant to describe whatever psychological processes are involved in coming to see an a priori proposition as evident. Both Bonjour (1998, p. 109) and Goldman (1999, p. 12) have noted that rational intuition could be naturalized along these lines.

in Section 3, the range of different kinds of beliefs, inference rules, or claims that we might classify as a priori may not admit of a single positive characterization. If so, there may be multiple forms of a priori justification, depending on the target class. Second, different standards apply to different forms of positive epistemic appraisal: the conditions that must be met for entitlement or epistemically-blameless reasoning are weaker than what must be met for justification sufficient for knowledge, and furthermore, there is disagreement with respect to what the standards for knowledge must be (although a naturalist will likely deny that they are the standards of the skeptic). One would expect a priori standards to vary concomitantly. Finally, the term ‘justification’ may be used to encompass a plurality of epistemic values, each of which may have different satisfaction conditions, and no one of which is equivalent to justification itself. For example, judgments with respect to the utility of a belief-forming process or method need not invoke an access condition that must be met by an epistemic agent, whereas a judgment that a belief is formed in a responsible manner will require that the agent be guided, in at least some sense, by appropriate considerations. If this is right, which concerns are being addressed will determine what conditions have to be met for justification.

Although it might be argued that some conceptual truths are analytic, it is generally conceded that other a priori subject matters, for example logic, can’t be explained in terms of analyticity. A number of different explanations of how logic could be a priori-justified that don’t appeal to rational intuition have been advanced by contemporary philosophers: Paul Boghossian (2000; 2001; 2003) and Christopher Peacocke (1993; 1998) have advocated meaning-based strategies, according to which the meaning-constituting roles of logical principles assure their validity; Boghossian also argues for a rule-circular justification for logic, as does Michael Dummett (1973).

Some philosophers have argued more broadly for a naturalistic account of a priori justification or knowledge: Louise Antony (2004), Alvin Goldman (1999), and Georges Rey (1998) defend a reliabilist account, according to which some reliable cognitive processes could be a priori warranters. Stephen Schiffer (unpub. ms.) identifies a way in which we could be non-inferentially justified in believing certain necessary truths based on how we are built as information processors. Hartry Field (1998; 2000) argues that basic belief-forming methodologies are default reasonable but empirically indefeasible, where these attributes are evaluative judgments that we make. One popular line of argumentation locates the source of entitlement or justification for basic inference rules in our need to presuppose them in reasoning. This approach is developed in different ways by Ernest Nagel (1956), Thomas Nagel (1997) and Crispin Wright (2004), among others. Questions may be raised about the adequacy of these various accounts, but they are naturalistic alternatives to the traditional appeal to rational insight.

## **6. The Philosophical Role of a Naturalistic A Priori**

In this final section, I consider some objections to a naturalistic conception of the a priori. For one thing, it might be thought that the naturalistic notion of the a priori that I have defended here is of little interest. The features of the a priori that rendered it suitable for a foundationalist account of knowledge – certainty, rational irrevisability, and the link with necessary truth – have either been rejected

or deemed inessential to the basic concept. However, empirical indefeasibility, albeit on a fallibilist interpretation, has been retained. So if there are claims, rules, or principles that have such a priori justification conditions, this is of considerable interest and epistemological importance. Moreover, adopting a naturalistic conception of the a priori leaves open the possibility that there may be a close connection between a priori justification and necessary truth: naturalistic apriorism need only deny that any apriori-justified proposition is thereby metaphysically necessary. But a naturalist can allow that the a priori may be intimately linked with necessary truth on some reasonably-naturalistic understanding of modal notions.<sup>31</sup>

A second worry is that any naturalistic form of a priori justification would have to be psychologistic and that, therefore, it would fail to establish the objectivity of a target claim. For example, if claims are allegedly justified by their self-evidence or the inconceivability of their falsehood, the worry is that their verisimilitude does not guarantee their truth (or in the case of inference rules, their seeming to be reliable doesn't guarantee their actual reliability). However, reliabilist accounts of a priori knowledge are paradigmatically naturalistic, and although they are psychologistic in the sense that they locate warrant within the psychological processes and methods that generate belief, if correct, they cannot be faulted for failing to establish the objectivity of their target beliefs: a belief is apriori-justified on a reliabilist account of justification just in case it is solely the result of a process that is, in fact, reliable and is indifferent to the particulars of experiential input. On a more internalistic account of justification, objectivity might be secured by an appeal to the overall coherence of a priori claims within a system of beliefs. It is open to the apriorist as well as to the empiricist to argue for the objectivity of traditionally a priori subject matters – such as mathematics and logic – by virtue of their instrumentality to the acquisition of knowledge generally, including their role in establishing scientific claims. Logic and mathematics are instrumental to science, so they “share in the success of science”; however, logic and mathematics are not empirical if they are *presupposed* rather than tested by scientific methods.

This brings me to a third worry, which is that fallibilist apriorism is indistinguishable from Quinean empiricism. It should be clear from what has been said that this is not the case. However, the point merits some further discussion because it is not as widely appreciated as it should be that fallibilism about the a priori is not equivalent to empiricism. Fallibilism allows that we don't know that a claim (or a warrant for a claim) will not come to be subsequently overturned on empirical grounds when future conceptual advancements and observations we might make in light of them are taken into consideration. But acknowledging this purely epistemic possibility does not tell us what basis we have for accepting the claim to begin with or what its justification conditions are. By contrast, Quine's revisability thesis is intended to show that no statement has

---

<sup>31</sup> Establishing the claim that a priori and necessary propositions are coextensive may be a challenge: while it is widely recognized that there are contingently true a priori statements that acquire their a priori status in virtue of choices we make about fixing the references of our terms, Tim Williamson (1986) and John Hawthorne (2002) have argued, respectively, that there either are or may be deeply contingent a priori statements for which there is no semantic guarantee that some verifying state of affairs actually exists.

the status of being a priori.<sup>32</sup> In order to do this, the revisability claim must be interpreted in a way that is stronger than mere fallibilism.

Quine's revisability claim must be understood as the claim that it is genuinely – not just purely epistemically – possible that P should be revised in light of experience. But this means that there are possible observations we could make such that if we did make them, revisions to an allegedly apriori-justified claim, for example a logical principle, would be rationally mandated or at least rational to accept. On the Quinean picture, a revision would be rational if making the revision would maximize simplicity and conservativeness in our overall system of belief. So Quine would appear to be committed to the following:

**REV:** For any given logical principle L (where logical principles should be understood to include valid rules of inference), there are possible observations we could make such that a revision to L would best preserve simplicity and conservativeness in our system of beliefs.

There are challenges to upholding REV. For one thing, REV requires that there be possible observations that would justifiably undermine L. It is hard to think of convincing cases of empirical observations that should count as evidence against elementary principles such as the law of identity or modus ponens. However, given the peculiarities of quantum mechanics, perhaps there are such possible observations. I think the harder problem to address is the claim that *a revision* to any given logical principle L would best preserve simplicity and conservativeness for some, but not all, possible observations that we might make. (If a revision would best preserve simplicity and conservativeness for any possible observations that we might make, then it would do so independently of the evidence, in which case the revision would not be made on empirical grounds.)

A revision of a logical principle involves more than just the abrogation of a particular principle in a particular domain of discourse; a revision requires the adoption of an alternative logic that does without the impugned principle. Consider once again the proposal to revise the distributive law. A quantum logic that does without the distributive law might be developed as a “special-purpose” logic to model the behavior of subatomic particles, but if this quantum logic is adopted as a revision to our general-purpose logic, the distributive law will no longer have the status of a logical principle.<sup>33</sup> Presumably we would still want to

---

<sup>32</sup> This general point is emphasized by both Hartry Field (1998; 2000) and Georges Rey (1993; 2005). Alternatively, if one were to accept Quine's extreme version of confirmation holism advanced in “Two Dogmas of Empiricism,” all claims would automatically have empirical justification conditions because the web is seamless and the system as a whole is the unit of confirmation. But this extreme version of confirmation holism is implausible, and Quine himself rejects it in *Word and Object* (1960, p. 13, fn).

<sup>33</sup> A nonclassical logic might be developed for the purpose of modeling an area of discourse that is not immediately and directly amenable to representation in classical logic, either because some class of statements violates an initial idealization assumption of classical logic or because an apparent violation of a logical principle is discovered within a particular subject matter. Free logics, logics of vagueness, and intuitionist logics for mathematics would be examples of the former, and quantum logics would be examples of the latter. Whether a “special-purpose” logic should count as a revision to our logic may depend on whether it also is being advanced as a general-purpose logic.

use the distributive laws when reasoning about macroscopic phenomena, so conditions under which the law could be employed would need to be specified. The proposed quantum logic would have to be such that the resultant gain in simplicity and conservativeness of adopting the new logic more than offset the concomitant loss of the distributive law as a logical principle. Perhaps the quantum logic would be judged to meet this goal. But the claim of REV is that *for each logical principle L*, there is a revision in light of possible observations we might make such that the revision would best preserve the dual goals of simplicity and conservativeness in the system. This is a very strong claim, and it is not one for which the vague picture of the web of belief offers any support.

## 7. Concluding Remarks

Debates about a priori justification are often characterized as debates between the apriorist and the naturalist. I hope to have shown that this is a mistake. A suitably-modest and fallibilist notion of a priori justification is, in principle, compatible with naturalism – unless naturalism is partially *defined* as the rejection of a priori knowledge. However, if the a priori is divested of properties that are extraneous to the basic a priori/a posteriori distinction, the motivation to equate naturalism with empiricism is largely eliminated. One can allow that there may be a priori forms of justification, yet be a thorough-going naturalist. The interesting question is whether any beliefs, claims, or rules have a priori justification conditions, and if so, what form that justification would take.

**Acknowledgments** This paper began as a talk for SUNY Albany's Philosophy Department, which I gave in May 2006. I subsequently presented versions of the paper at the British Society for Philosophy of Science Annual Conference and the Joint Session of the Mind Association and Aristotelian Society in July 2007, and to the University of Connecticut Philosophy Department in October 2007. I am grateful to audience members at all the talks, especially JC Beall, Tom Bontly, Otávio Bueno, Alexander Jackson, Michael Lynch, Robert Meyers, and Ron McClamrock, for helpful questions and comments. I would like to thank Ken Akiba, Michael Bishop and an anonymous referee from this journal for helpful comments on a draft of the paper, and Brad Armour-Garb for stimulating discussion. Special thanks are due to Jonathan Adler and Hartry Field for many helpful comments, both in response to earlier drafts of the paper and in conversation.

## References

- Adler, J. (2002). *Belief's Own Ethics*. (Cambridge, MA: MIT Press).
- Antony, L. (2004). A Naturalized Approach to the A Priori. *Philosophical Issues*, 14, *Epistemology*, 1-17.
- Adams, E.W. (1975.) *The Logic of Conditionals: An Application of Probability to Deductive Logic*. (Dordrecht: D. Reidel).
- Audi, R. (2000). Philosophical Naturalism at the Turn of the Century. *Journal of Philosophical Research*, 25, 27-45.
- Bealer, G. (2002). Modal Epistemology and the Rationalist Renaissance. (In J. Hawthorne and T. Szabó-Gendler (Eds.), *Conceivability and Possibility* (pp. 71-126). New York: Oxford.)
- Boghossian, P. (2000). Knowledge of Logic. (In P. Boghossian and C. Peacocke (Eds.), *New Essays on the A Priori* (pp. 229-54). Oxford: Oxford University Press).
- Boghossian, P. (2001). How Are Objective Epistemic Reasons Possible? *Philosophical Studies*, 106, 1-40.

- Boghossian, Paul. (2003). Blind Reasoning. *Proceedings of the Aristotelian Society, Supplementary Volume*, 77 (1), 225-248.
- Bonjour, L. (1998). *In Defense of Pure Reason*. (Cambridge: Cambridge University Press).
- Burge, T. (1993). Content Preservation. *Philosophical Review*, 102, 457-488.
- Burge, T. (1998). Computer Proof, A Priori Knowledge, and Other Minds. (In J. Tomberlin (Ed.), *Philosophical Perspectives* 12, Language, Mind, and Ontology. Malden, MA: Blackwell.)
- Chisholm, R. (1977). *Theory of Knowledge*. (Englewood, NJ: Prentice Hall).
- Casullo, A. (2003). *A Priori Justification*. (New York: Oxford University Press).
- DeCaro, M. And Macarthur, D. (Eds.) (2004). *Naturalism in Question*. (Cambridge, MA and London: Harvard University Press).
- DeRose, K. (1991). Epistemic Possibilities. *Philosophical Review*, 100, 581-605.
- DeRose, K. (1998) Simple 'Might's', Indicative Possibilities and the Open Future. *Philosophical Quarterly*, 48, No. 190, 67-82.
- Devitt, M. (1996). *Coming to Our Senses: A Naturalistic Program for Semantic Localism*. (Cambridge: Cambridge University Press).
- Dummett, M. (1973). The Justification of Deduction. Lecture. British Academy, London. In M. Dummett, *Truth and Other Enigmas* (pp. 290-318). Cambridge, MA: Harvard University Press).
- Egan, A. (2007). Epistemic Modals, Relativism, and Assertion. *Philosophical Studies*, 133 (1), 1-22.
- Egan, A, Hawthorne, J., and Weatherson, B. Epistemic Modals in Context. In G. Preyer and G. Peter (Eds.), *Contextualism in Philosophy* (pp. 131-168), Oxford: Oxford University Press.
- Evnine, S. (2001). Learning from One's Mistakes: Epistemic Modality and the Nature of Belief. *Pacific Philosophical Quarterly*, 82, 157-77.
- Ewald, W. (Ed.) (1996). *From Kant to Hilbert*, Vol. I. (Oxford: Oxford University Press).
- Field, H. (1998). Epistemological Nonfactualism and the A Prioricity of Logic. *Philosophical Studies*, 92(1-2), 1-24.
- Field, H. (2000). A Prioricity as an Evaluative Notion. (In P. Boghossian and C. Peacocke (Eds.), *New Essays on the A Priori* (pp. 117-49), Oxford: Oxford University Press; and in H. Field, *Truth and the Absence of Fact*, (pp. 361-90), New York: Oxford Univ. Press.)
- Firth, R. (1967). The Anatomy of Certainty. *Philosophical Review*, Vol. 76, No. 1, 3-27.
- Goldman, A. (1994). Naturalistic Epistemology and Reliabilism. (In P. French, T. Uehling, H. Wettstein (Eds.), *Midwest Studies in Philosophy*, 19, *Philosophical Naturalism*, (pp. 301-20). Notre Dame, IN: Univ. of Notre Dame Press.)
- Goldman, A. (1999). A Priori Warrant and Naturalistic Epistemology. (In J. Tomberlin, (Ed.), *Philosophical Perspectives*, 13, *Epistemology*, (pp. 1-28). Cambridge, MA; Oxford: Blackwell.)
- Haack, S. (1979). Fallibilism and Necessity. *Synthese*, 41, 37-63.
- Hacking, I. (1967). Possibility. *Philosophical Review*, 76, no. 2, 143-68.
- Hacking, I. (1975). All Kinds of Possibility. *Philosophical Review*, 84, 321-37.
- Hawthorne, J. (2002). Deeply Contingent A Priori Knowledge. *Philosophy and Phenomenological Research*, 65, No. 2, 247-269.
- Hookway, C. (2007). Fallibilism and the Aim of Inquiry. *Aristotelian Society Supplementary Volume* 81, 1-22.
- Kitcher, P. (1992). The Naturalist Returns. *Philosophical Review*, 101, No. 1, *Philosophy in Review: Essays on Contemporary Philosophy*. Jan., 1992, 53-114.
- Kripke, S. (1972/80). *Naming and Necessity*. (Cambridge, MA: Harvard University Press).
- Lehrer, K. and Kim, K. (1990). The Fallibility Paradox. *Philosophy and Phenomenological Research*, 50, Supplement, 99-107.

- Lycan, W. (1993). MPP, Rip. In J. Tomberlin (Ed.), *Philosophical Perspectives*, 7, *Language and Logic*, 411-428.
- MacFarlane, J. (2005a). The Assessment Sensitivity of Knowledge Attributions. (In Hawthorne, J. and T. Szabó-Gendler (Eds.), *Oxford Studies in Epistemology*, 1, (pp. 197-253). Oxford: Oxford University Press.)
- MacFarlane, J. (2005b). Making Sense of Relative Truth. *Proceedings of the Aristotelian Society*, 105, 321-339.
- Maddy, P. (2000). Naturalism and the A Priori. (In P. Boghossian and C Peacocke (Eds.), *New Essays on the A Priori* (pp. 92-116). Oxford: Oxford University Press.)
- Maffie, J. (1990). Recent Work on Naturalized Epistemology. *American Philosophical Quarterly*, 27, No. 4, 281-292.
- McGee, V. (1985). A Counterexample to Modus Ponens. *Journal of Philosophy*, 82, 462-471.
- Meyers, R. (1988). *The Likelihood of Knowledge*. (Dordrecht: Kluwer).
- Moore, G. E. (1962). *Commonplace Book*, 1919-53, H. D. Lewis, (Ed.), (London: George Allen & Unwin).
- Nagel, E. (1956). *Logic Without Metaphysics*. (Glencoe, IL: Free Press).
- Nagel, T. (1997). *The Last Word*. (New York: Oxford University Press).
- Peacocke, C. (1993). How Are A Priori Truths Possible? *European Journal of Philosophy*, 1, No. 2, 175-199.
- Peacocke, C. (1998). Implicit Conceptions, Understanding and Rationality. (In E. Villanueva (Ed.), *Philosophical Issues*, 9, *Concepts*, (pp. 43-87). Atascadero, CA: Ridgeview Publishing.)
- Peirce, C.S. (1931-1958). *Collected Papers*. C. Hartshorne, P. Weiss, and A.S. Burks (Eds.) (Cambridge, MA).
- Quine, W. V. O. (1953). Two Dogmas of Empiricism. (In W.V.O. Quine, *From a Logical Point of View*. 2<sup>nd</sup> Edition, (pp. 20-46). Cambridge, MA: Harvard University Press.)
- Quine, W.V.O. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Quine, W. V. O. (1981). *Theories and Things*. (Cambridge, MA: MIT Press).
- Rey, G. (1993). The Unavailability of What We Mean I: A Reply to Quine, Fodor and LePore. *Grazer Philosophical Studien*, 96, 61-101. (Reprinted in J. Fodor (Ed.), *Holism: A Consumer Update*. Amsterdam: Rodopi.)
- Rey, G. (1998). A Naturalistic A Priori. *Philosophical Studies*, 92, 25-43.
- Rey, G. (2005). The Rashness of Traditional Rationalism and Empiricism. In M. Ezcúrdia, R. Stainton, and C. Viger (Eds.), *New Essays in the Philosophy of Language and Mind*, *Canadian Journal of Philosophy*, Supplementary Volume 30, 227-58.
- Richard, M. (2004). Contextualism and Relativism. *Philosophical Studies*, 119, 214-242.
- Russell, B. (1948). *Human Knowledge*. (New York: Simon Schuster).
- Siegel, H. (1995). Naturalized Epistemology and 'First Philosophy'. *Metaphilosophy*, 26, 46-62.
- Stroud, B. (1996). The Charm of Naturalism. *Proceedings and Addresses of the American Philosophical Association*, 70, no. 2. Newark, DE: American Philosophical Association. (Reprinted in M. De Caro and D. MacArthur (Eds.), *Naturalism in Question*. Cambridge, MA and London: Harvard University Press.)
- Szabó-Gendler, T. and Hawthorne, J. (2002). *Conceivability and Possibility*. (Oxford: Clarendon Press).
- Teller, P. (1972). Epistemic Possibility. *Philosophia*, 2, 303-320.
- Williamson, T. (1986). The Contingent A Priori: Has It Anything to Do with Indexicals? *Analysis*, 113-117.
- Wright, C. (2004). Warrant for Nothing: Notes on Epistemic Entitlement. *Proceedings of the Aristotelian Society*, Supplementary Volume 78, 167-212.
- Yablo, S. (1993). Is Conceivability a Guide to Possibility? *Philosophy and Phenomenological Research*, 53, No. 1, 1-42.
- Yablo, S. (2002). Coulda, Woulda, Shoulda. (In J. Hawthorne and T. Szabó-Gendler (Eds.), *Conceivability and Possibility*, (pp. 441-492). New York: Oxford.)