How Bayesian Confirmation Theory Handles the Paradox of the Ravens

Branden Fitelson

University of California–Berkeley

branden@fitelson.org

James Hawthorne

University of Oklahoma

hawthorne@ou.edu

1. Introduction

The Paradox of the Ravens (*a.k.a,*, The Paradox of Confirmation) is indeed an old chestnut. A great many things have been written and said about this paradox and its implications for the logic of evidential support.[1] The first part of this paper will provide a brief survey of the early history of the paradox. This will include the original formulation of the paradox and the early responses of Hempel, Goodman, and Quine. The second part of the paper will describe attempts to resolve the paradox within a Bayesian framework, and show how to improve upon them. This part begins with a discussion of how probabilistic methods can help to clarify the statement of the paradox itself. And it describes some of the early responses to probabilistic explications. We then inspect the assumptions employed by traditional (canonical) Bayesian approaches to the paradox. These assumptions may appear to be overly strong. So, drawing on weaker

[1] For a nice taste of this voluminous literature, see the bibliography in (Vranas 2004).

assumptions, we formulate a new-and-improved Bayesian confirmation-theoretic resolution of the Paradox of the Ravens.

2. The Original Formulation of the Paradox

Traditionally, the Paradox of the Ravens is generated by the following two assumptions (or premises).

- **Nicod Condition (NC)**: For any object *a* and any predicate *F* and *G*, the proposition that *a* has both *F* and *G* confirms the proposition that every *F* has *G*. A more formal version of (NC) is the following claim: $(Fa \cdot Ga)$ confirms $(\forall x)(Fx \supset Gx)$, for any individual term '*a*' and any pair of predicates '*F*' and '*G*'.
- **Equivalence Condition (EC)**: For any propositions H_1, H_2, and E, if E confirms H_1 and H_1 is (classically) logically equivalent to H_2, then E confirms H_2.

From (NC) and (EC), we can deduce the following, "paradoxical conclusion":

- **Paradoxical Conclusion (PC)**: The proposition that *a* is both non-black and a non-raven, $({\sim}Ba \cdot {\sim}Ra)$, confirms the proposition that every raven is black, $(\forall x)(Rx \supset Bx)$.

The canonical derivation of (PC) from (EC) and (NC) proceeds as follows:

1. By (NC), $({\sim}Ba \cdot {\sim}Ra)$ confirms $(\forall x)({\sim}Bx \supset {\sim}Rx)$.
2. By Classical Logic, $(\forall x)({\sim}Bx \supset {\sim}Rx)$ is equivalent to $(\forall x)(Rx \supset Bx)$.
3. By (1), (2), and (EC), $({\sim}Ba \cdot {\sim}Ra)$ confirms $(\forall x)(Rx \supset Bx)$. *QED*.

The earliest analyses of this infamous paradox were offered by Hempel, Goodman, and Quine. Let's take a look at how each of these famous philosophers attempted to resolve it.

3. Early Analyses of the Paradox due to Hempel, Goodman, and Quine

3.1 The Analyses of Hempel and Goodman

Hempel (1945) and Goodman (1954) didn't view (PC) as paradoxical. Indeed, Hempel and Goodman viewed the argument above from (1) and (2) to (PC) as *sound*. So, as far as Hempel and Goodman are concerned, there is something misguided about whatever intuitions may have lead some philosophers to see "paradox" here. As Hempel explains (Goodman's discussion is very similar on this score), one might be misled into thinking that (PC) is false by conflating (PC) with a different claim (PC*) – a claim that is, in fact, false. Hempel warns us that [our emphasis]

> …in the seemingly paradoxical cases of confirmation, we are often not judging the relation of the given evidence *E alone* to the hypothesis *H* … instead, we tacitly introduce a comparison of *H* with a body of evidence which consists of *E* in conjunction with an additional amount of information we happen to have at our disposal.

We will postpone discussion of this crucial remark of Hempel's until the later sections on Bayesian clarifications of the paradox – where its meaning and significance will become clearer. Meanwhile, it is important to note that Hempel and Goodman also provide independent motivation for premise (1) of the canonical derivation of (PC) – a motivation independent of (NC) – in an attempt to further bolster the traditional argument.[2] The following argument for premise (1) is endorsed by both Hempel and Goodman [our emphasis and brackets]:

---

[2] Almost all early commentators on the paradox have viewed (EC) and premise (2) as beyond reproach. But not all contemporary commentators are so sanguine about (EC) and (2). See (Sylvan and Nola 1991) for detailed discussion of non-classical logics and the paradoxes of confirmation. See (Gemes 1999) for a probabilistic approach that also denies premise (2). We will not discuss such approaches here. We restrict our focus to accounts couched in terms of classical logic.

> If the evidence E consists *only* of one object which … is a non-raven [$\sim Ra$], then E may reasonably be said to confirm that all objects are non-ravens [$(\forall x) \sim Rx$], and *a fortiori*, E supports the weaker assertion that all non-black objects are non-ravens [$(\forall x)(\sim Bx \supset \sim Rx)$].

This alternative argument for premise (1) presupposes the *Special Consequence Condition*:

> (SCC) For all propositions H_1, H_2, and E, if E confirms H_1, and H_1 (classically) logically entails H_2, then E confirms H_2.

Early instantial and probabilistic theories of confirmation [*e.g.*, those presupposed by Hempel, Goodman, Quine, Carnap (1950)] embraced (SCC). But, from the point of view of contemporary Bayesian confirmation theory, (SCC) is false, as was first shown by Carnap (1950).[3] We will return to this recent dialectic below, in our discussion of the paradox within the context of contemporary Bayesian confirmation theory. But before making the transition to Bayesian confirmation, let us briefly discuss Quine's rather influential response to the paradox, which deviates significantly from the views of Hempel and Goodman.

3.2 Quine on the Paradox of the Ravens

In his influential paper "Natural Kinds", Quine (1969) offers an analysis of the paradox of confirmation that deviates radically from the Hempel-Goodman line. Unlike Hempel and Goodman, Quine *rejects* the paradoxical conclusion (PC). Since Quine accepts classical logic,

[3] In response to Karl Popper's (1954) critique of *Logical Foundations of Probability*, Carnap acknowledges in the introduction to the second edition (1962) that he had two distinct notions of confirmation (i.e. high probability above a threshold, and increase in probability) in mind when he wrote different parts of the book. This explains how he could apparently accept (SCC) at one place and reject it at a different place in the same book.

this forces him to reject either premise (1) or premise (2) of the (classically valid) canonical argument for (PC). Since Quine also accepts the (classical) equivalence condition (EC), he must accept premise (2). Thus, he is led, inevitably, to the rejection of premise (1). This means he must reject (NC) – and he does so. Indeed, according to Quine, not only does $({\sim}Ba \cdot {\sim}Ra)$ *fail to confirm* $(\forall x)({\sim}Bx \supset {\sim}Rx)$, but also ${\sim}Ra$ *fails to confirm* $(\forall x){\sim}Rx$. According to Quine, the failure of instantial confirmation in these cases stems from the fact that the predicates 'non-black' [${\sim}B$] and 'non-raven' [${\sim}R$] are not *natural kinds* – i.e., the objects falling under ${\sim}B$ and ${\sim}R$ are not sufficiently similar to warrant instantial confirmation of universal laws involving ${\sim}B$ or ${\sim}R$. Only instances falling under natural kinds can warrant instantial confirmation of universal laws. Thus, for Quine the source of the problem is (NC). He suggests that the *unrestricted* version of (NC) is false, and must be replaced by a *restricted* version that applies only to natural kinds:

> **Quine–Nicod Condition (QNC)**: For any object a and any *natural kinds* F and G, the proposition that a has both F and G confirms the proposition that every F has G. More formally, $(Fa \cdot Ga)$ confirms $(\forall x)(Fx \supset Gx)$, for any individual term a, provided that the predicates 'F' and 'G' refer to *natural kinds*.

To summarize, Quine thinks (PC) is false, and that the (valid) canonical argument for (PC) is unsound because (NC) is false. Furthermore, according to Quine, once (NC) is restricted in scope to *natural kinds*, the resulting restricted instantial confirmation principle (QNC) is true, but useless for deducing (PC).[4] However, many other commentators have taken (NC) to be the real

[4] Interestingly, while Hempel and Goodman are completely unsympathetic to Quine's strategy *here*, they are much more sympathetic to such maneuvers in the context of the *Grue* Paradox. In this sense, Quine's approach to the paradoxes is more unified and systematic than Hempel's or Goodman's, since they give "special treatment" to Grue-

culprit here, as we'll soon see. We think that the real problems with (NC) [and (QNC)!] only become clear when the paradox is cast in more precise Bayesian terms, in a way that will be explicated in the second part of this paper. But we will first show how the Bayesian framework allows us to clarify the paradox and the historical debates surrounding it.

4. Bayesian Clarifications of (NC) and (PC)

Hempel (1945) provided a cautionary remark about the paradox. He warned us not to conflate the paradoxical conclusion (PC) with a distinct (intuitively) *false* conclusion (PC*) that (intuitively) does *not* follow from (NC) and (EC). We think Hempel's intuitive contrast between (PC) and (PC*) is important for a proper understanding the paradox. So, we'll discuss it briefly.

What, precisely, was the content of this (PC*)? Well, that turns out to be a bit difficult to say from the perspective of traditional, deductive accounts of confirmation. Based on the rest of Hempel's discussion and the penetrating recent exegesis of Patrick Maher (Maher 1999), we think the most accurate informal way to characterize (PC*) is as follows:

> (PC*) If one observes that an object *a* – *already known to be a non-raven* – is non-black (hence, is a non-black non-raven), then *this* observation confirms that all ravens are black.

As Maher points out, it is somewhat tempting to conflate (PC*) and (PC). But, Hempel did *not* believe that (PC*) was true (intuitively) about confirmation, nor did he think that (PC*)

---

predicates, while Quine views the problem – in *both* paradoxes of confirmation – to be rooted in the "non-naturalness" of the referents of the predicates involved. For what it's worth, we think a unified and systematic approach to the paradoxes is to be preferred. But, we think a unified *Bayesian* approach is preferable to Quine's instantial approach. However, our preferred Bayesian treatment of Grue will have to wait for another paper.

(intuitively) follows from (NC) and (EC). This is because, intuitively, *observing (known) non-ravens does not tell us anything about the color of ravens*. While this seems intuitively quite plausible, it is quite difficult to see how Hempel's *confirmation theory* can *theoretically ground* the desired distinction between (PC) and (PC*). What Hempel *says* is that we should not look at the evidence E in *conjunction* with other information that we might have at our disposal. Rather, we should look at the confirmational impact of learning *E and only E*.

There are two problems with this (the second worse than the first). First, as we have cast it (and as we think it *should* be cast), (PC*) is not a claim about the confirmational impact on $(\forall x)(Rx \supset Bx)$ of learning $\sim Ba$ in *conjunction* with other information about a (*i.e.*, $\sim Ra$), but the impact on $(\forall x)(Rx \supset Bx)$ of learning $\sim Ba$ *given that you already know* $\sim Ra$. Basically, we are distinguishing the following two kinds of claims:

- E confirms H, *given* A – *e.g.*, $\sim Ba$ confirms $(\forall x)(Rx \supset Bx)$, *given* $\sim Ra$ – *versus*
- $(E \cdot A)$ confirms H, *unconditionally* – *e.g.*, $(\sim Ba \cdot \sim Ra)$ confirms $(\forall x)(Rx \supset Bx)$ *unconditionally*.

Note: in classical deductive logic, there *is no distinction* between:

- X entails Y, *given* Z, and
- $(X \cdot Z)$ entails Y.

For this reason, Hempel's theory of confirmation (which is based on deductive entailment – see below) is incapable of making such a distinction. Perhaps this explains why he states things in terms of *conjunction*, rather than *conditionalization*. After all, he offers no confirmation-theoretical distinction between 'and' and 'given that'. So, while it seems that there is an *intuitive* distinction of the desired kind between (PC) and (PC*), it is unclear how Hempel's *theory* is supposed to make this distinction formally precise [see Maher (1999) for discussion].[5]

The second problem with Hempel's intuitive "explaining away" of the paradox is far more worrisome. As it turns out, Hempel's official theory of confirmation is *logically incompatible* with his intuitive characterization of what is going on. According to Hempel's theory of confirmation, the confirmation relation is *monotonic*. That is, Hempel's theory entails:

> (M) If *E* confirms *H*, relative to *no* (or tautological) background information, then *E* confirms *H* relative to *any* collection of background information whatsoever.

The reason Hempel's theory entails (M) is that it explicates "*E* confirms *H* relative to *K*" as "*E* & *K* entails *X*", where the proposition *X* is obtained from the syntax of *H* and *E* in a certain complex way, which Hempel specifies (the technical details of Hempel's approach to confirmation won't matter for present purposes). Of course, if *E by itself* entails *X*, then so does *E* & *K*, for any *K*.[6] Thus, according to Hempel's theory of confirmation, if (PC) is true, then

---

[5] Perhaps Hempel had something like the following in mind. Notice that $(\forall x)(Rx \supset Bx)$ entails *Ba* given *Ra*; so, given *Ra*, ~*Ba* falsifies $(\forall x)(Rx \supset Bx)$ and, on Hempel's account, *Ba* confirms it. Likewise, $(\forall x)(Rx \supset Bx)$ entails ~*Ra* given ~*Ba*; so, given ~*Ba*, *Ra* falsifies $(\forall x)(Rx \supset Bx)$ and, on Hempel's account, ~*Ra* confirms it. However, $(\forall x)(Rx \supset Bx)$ entails neither *Ba* nor ~*Ba* given ~*Ra*; so, arguably, one might hold that $(\forall x)(Rx \supset Bx)$ cannot be confirmed by either *Ba* or by ~*Ba* given ~*Ra* (though, as already affirmed, it is confirmed by ~*Ra* given ~*Ba*). Similarly, $(\forall x)(Rx \supset Bx)$ entails neither *Ra* nor ~*Ra* given *Ba*; so, arguably, one might hold that $(\forall x)(Rx \supset Bx)$ cannot be confirmed by either *Ra* or by ~*Ra* given *Ba* (though, of course, it is confirmed by *Ba* given *Ra*). Even if a Hempelian story along these lines can be told, it won't save Hempel's analysis from problem #2, below.

[6] Hypothetico-deductive approaches to confirmation also imply (M), since they explicate "*E* confirms *H* relative to *K*" as "*H* & *K* entails *E*." So, H-D confirmation cannot avail itself of a Hempel-style resolution of the paradox either.

(PC*) must also be true. So, while intuitively compelling and explanatory, Hempel's suggestion that (PC) is true but (PC*) is false *contradicts* his own theory of confirmation. As far as we know, this logical inconsistency in Hempel (and Goodman's) discussions of the paradox of confirmation has not been discussed in the literature.[7]

It is clear that Hempel was onto something important here with his intuitive distinction between claims (PC) and (PC*), but his confirmation theory just lacked the resources to properly spell out his intuitions. Here contemporary Bayesian confirmation theory really comes in handy.

According to Bayesian confirmation theory, "E confirms H, given K", and "$(E \cdot K)$ confirms H, *unconditionally*" have quite different meanings. Essentially, this is possible because Bayesian explications of the confirmation relation do *not* entail *monotonicity* (M). Specifically, contemporary Bayesians offer the following account of conditional and unconditional confirmation – where hereafter, we will use the words "confirms" and "confirmation" in accordance with this Bayesian account:

[7] Maher notes that Hempel *never proves* that (PC) is true while (PC*) is false. This is an understatement! He *cannot* prove this claim, *on pain of contradiction with his official theory of confirmation*. We think the reason Hempel (and others) missed this inconsistency is that it is easy to conflate *objectual* and *propositional* senses of "confirmation". If you think of the *objects* doing the confirming, then one can see (PC) as true and (PC*) as false (even from a deductivist point of view). But, if you think of the *propositions* as doing the confirming, then this is impossible from a deductivist point of view (*i.e.*, from the point of view of any theory which entails (M)). The salient passages from Hempel suggest that he slides back and forth between objectual and propositional senses of confirmation. And, we suspect that this is what led him into the present inconsistency.

- **Bayesian Confirmation**. E confirms H, *given* K (or *relative to* K), just in case $P[H \mid E \cdot K] > P[H \mid K]$. And, E confirms H, *unconditionally*, just in case $P[H \mid E] > P[H]$, where P[•] is some suitable probability function.[8]

It is easy to see, on this account of (conditional and unconditional) confirmation, that there will be a natural distinction between (PC) and (PC*). From a Bayesian point of view this distinction becomes:

(PC) $P[(\forall x)(Rx \supset Bx) \mid {\sim}Ba{\cdot}{\sim}Ra] > P[(\forall x)(Rx \supset Bx)]$, and

(PC*) $P[(\forall x)(Rx \supset Bx) \mid {\sim}Ba{\cdot}{\sim}Ra] > P[(\forall x)(Rx \supset Bx) \mid {\sim}Ra]$

What Hempel had in mind (charitably) is the former, *not* the latter. This is crucial for understanding the ensuing historical dialectic regarding the paradox. The important point here is that Bayesian confirmation theory has the theoretical resources to distinguish conditional and unconditional confirmation, but traditional (classical) deductive accounts do not. As a result Bayesian theory allows us to precisely articulate Hempel's intuition concerning why people might (falsely) believe that the paradoxical conclusion (PC) is false by conflating it with (PC*).

A key insight of Bayesian confirmation theory is that it represents confirmation as a *three*-place relation between *evidence E*, *hypothesis H*, and *background corpus K*. From this perspective the traditional formulation of the paradox is *imprecise* in an important respect: *it leaves unclear which background corpus is presupposed* in the (NC) – and, as a result, also in

---

[8] We won't bother to discuss the competing axiomatizations and interpretations of probability. These details won't matter for our discussion. For simplicity we will just assume that P is some rational credence function, and that it satisfies an appropriate version of the standard, (Kolmogorov 1956) axioms. But these assumptions could be altered in various ways without affecting the main points we will make below.

the (PC). In other words, there is a *missing quantifier* in the traditional formulations of (NC) and (PC). Here are four possible precisifications of (NC) [the corresponding precisifications of (PC) should be obvious]:

- (NC_w) For any individual term '*a*' and any pair of predicates '*F*' and '*G*', there is some possible background K such that $(Fa \cdot Ga)$ confirms $(\forall x)(Fx \supset Gx)$, *given K*.
- (NC_α) Relative to our *actual* background corpus K_α, for any individual term '*a*' and any pair of predicates '*F*' and '*G*', $(Fa \cdot Ga)$ confirms $(\forall x)(Fx \supset Gx)$, *given* K_α.
- ($NC_\top$) Relative to *tautological* (or *a priori*) background corpus $K_\top$, for any individual term '*a*' and any pair of predicates '*F*' and '*G*', $(Fa \cdot Ga)$ confirms $(\forall x)(Fx \supset Gx)$, *given* $K_\top$.
- (NC_s) Relative to *any possible* background corpus *K*, for any individual term '*a*' and any pair of predicates '*F*' and '*G*', $(Fa \cdot Ga)$ confirms $(\forall x)(Fx \supset Gx)$, *given K*.

Which rendition of (NC) is the one Hempel and Goodman had in mind? Well, (NC_w) seems *too weak* to be of much use. There is bound to be *some* corpus with respect to which non-black non-ravens confirm 'All non-black things are non-ravens', but this corpus may not be very interesting (e.g., the corpus which contains '$(\sim Ba \cdot \sim Ra) \supset (\forall x)(\sim Bx \supset \sim Rx)$'!).

What about (NC_α)? Well, that depends. If we happen to (actually) *already know* that $\sim Ra$, then all bets are off as to whether $\sim Ba$ confirms $(\forall x)(\sim Bx \supset \sim Rx)$, relative to K_α (as Hempel suggests, and Maher makes precise). So, only a suitably restricted version of (NC_α) would satisfy Hempel's constraint. (We'll return to this issue, below.)

How about (NC_s)? This rendition is *too strong*. As we'll soon see, I.J. Good *demonstrated* that (NC_s) is *false* in a Bayesian framework.

What about ($NC_\top$)? As Maher (1999) skillfully explains, Hempel and Goodman (and Quine) have something much closer to ($NC_\top$) in mind. Originally, the question was whether learning *only* ($\sim Ba \cdot \sim Ra$) *and nothing else* confirms that all ravens are black. And, it seems natural to understand this in terms of confirmation relative to "tautological (or *a priori*) background". We will return to the notion of "tautological confirmation", and the (NC_α) *vs* ($NC_\top$) controversy, below. But, first, it is useful to discuss I.J. Good's knock-down counterexample to (NC_s), and his later (rather lame) attempt to formulate a counterexample to ($NC_\top$).

5. I.J. Good's Counterexample to (NC_s) and his "Counterexample" to ($NC_\top$)

Good (1967) asks us to consider the following example (we're paraphrasing here):

- Our background corpus K says that exactly one of the following hypotheses is true: (H) there are 100 black ravens, no non-black ravens, and 1 million other birds, or else ($\sim H$) there are 1,000 black ravens, 1 white raven, and 1 million other birds. And K also states that an object a is selected at random from all the birds. Given this background K, we have:

$$P[Ra \cdot Ba \mid (\forall x)(Rx \supset Bx) \cdot K] = \frac{100}{1000100} < P[Ra \cdot Ba \mid \sim(\forall x)(Rx \supset Bx) \cdot K] = \frac{1000}{1001000}$$

Hence, Good has described a background corpus K relative to which ($Ra \cdot Ba$) *dis*confirms $(\forall x)(Rx \supset Bx)$. This is sufficient to show that (NC_s) is false.

Hempel (1967) responded to Good by claiming that (NC_s) is *not* what he had in mind, since it smuggles too much "unkosher" (*a posteriori*) empirical knowledge into K. Hempel's challenge to Good was (again, charitably) to find a counterexample to ($NC_{\top}$). Good (1968) responded to Hempel's challenge with the following much less conclusive (rather *lame*, we think) "counterexample" to ($NC_{\top}$) [our brackets]:

> …imagine an infinitely intelligent newborn baby having built-in neural circuits enabling him to deal with formal logic, English syntax, and subjective probability. He might now argue, after defining a [raven] in detail, that it is initially extremely unlikely that there are any [ravens], and therefore that it is extremely likely that all [ravens] are black. … On the other hand, if there are [ravens], then there is a reasonable chance that they are a variety of colours. Therefore, if I were to discover that even a black [raven] exists I would consider [$(\forall x)(Rx \supset Bx)$] to be less probable than it was initially.

Needless to say, this "counterexample" to ($NC_{\top}$) is far from conclusive! To us it seems completely unconvincing [see Maher (1999) for a trenchant analysis of this example]. The problem here is that in order to give a rigorous and compelling counterexample to ($NC_{\top}$), one needs a *theory* of "tautological confirmation" – i.e. of "confirmation relative to tautological background". Good *doesn't have* such a theory (nor do most contemporary probabilists), which explains the lack of rigor and persuasiveness of "Good's Baby". However, Patrick Maher does have such an account; and he has applied it in his recent, neo-Carnapian, Bayesian analysis of the paradox of the ravens.

6. Maher's Neo-Carnapian Analysis of the Ravens Paradox

Carnap (1950, 1952, 1971, 1980) proposed various theories of "tautological confirmation" in terms of "logical probability". Recently Patrick Maher (1999, 2004) has brought a Carnapian approach to bear on the ravens paradox, with some very enlightening results. For our purposes it is useful to emphasize two consequences of Maher's neo-Carnapian, Bayesian analysis of the paradox. First, Maher shows that (PC*) is *false* on a neo-Carnapian theory of (Bayesian) confirmation. That is, if we take a suitable class of Carnapian probability functions $\mathrm{P}_c(\bullet \mid \bullet)$ – *e.g.*, those of Maher (1999) – as our "probabilities relative to tautological background", then we get the following result [see Maher (1999)]

- $\mathrm{P}_c[(\forall x)(Rx \supset Bx) \mid {\sim}Ba \cdot {\sim}Ra] = \mathrm{P}_c[(\forall x)(Rx \supset Bx) \mid {\sim}Ra]$

Intuitively, this says that observing the color of (known) non-ravens tells us nothing about the color of ravens, relative to tautological background corpus. This is a theoretical vindication of Hempel's intuitive claim that (PC*) is false – a vindication that is at best difficult to make out in Hempel's deductive theory of confirmation. But, all is not beer and skittles for Hempel.

More recently, Maher (2004) has convincingly argued [contrary to what he had previously argued in his (1999)] that, within a proper neo-Carnapian Bayesian framework, Hempel's (NC_{T}) is false, and so is its Quinean "restriction" ($\mathrm{QNC}_{\mathrm{T}}$). That is, Maher (2004) has shown that (from a Bayesian point of view) *pace* Hempel, Goodman, *and* Quine, *even relative to tautological background, positive instances do not necessarily confirm universal generalizations – not even for generalizations that involve only natural kinds*! The details of Maher's counterexample to ($\mathrm{QNC}_{\mathrm{T}}$) [hence, to ($\mathrm{NC}_{\mathrm{T}}$) as well] would take us too far afield. But, we mention it here because it shows that probabilistic approaches to confirmation are much richer and more powerful than traditional, deductive approaches. And, we think, Maher's work *finally* answers Hempel's challenge to Good – a challenge that went unanswered for nearly forty years.

Moreover, Maher's results also suggest that Quine's analysis in "Natural Kinds" was off the mark. Contrary to what Quine suggests, the problem with (NC) is not merely that it needs to be restricted in scope to certain kinds of properties. The problems with (NC) run much deeper than that. Even the most promising Hempelian precisification of (NC) is false, and a restriction to "natural kinds" does not help (since Maher-style, neo-Carnapian counterexamples can be generated that employ only to "natural kinds" in Quine's sense).[9]

While Maher's neo-Carnapian analysis is very illuminating, it is by no means in the mainstream of contemporary Bayesian thought. Most contemporary Bayesians reject Carnapian logical probabilities and the Carnapian assumption that there is any such thing as "degree of confirmation relative to tautological background." Since contemporary Bayesians have largely rejected this project, they take a rather different tack to handle *the ravens* paradox.

7. The Canonical Contemporary Bayesian Approaches to the Paradox

Perhaps somewhat surprisingly, almost all contemporary Bayesians implicitly assume that the paradoxical conclusion is *true*. They hold that (PC) is not really paradoxical after all once we take into account the fact that confirmation really comes in degrees – that some evidence is more strongly confirmatory that other evidence. Thus, contemporary Bayesians aim to soften the impact of (PC) by establishing certain *comparative* and/or *quantitative* confirmational claims. Specifically, Bayesians typically aim to show (at least) that the observation of a black raven, $(Ba \cdot Ra)$, confirms "all ravens are black" *more strongly than* the observation of a non-black non-raven, $({\sim}Ba \cdot {\sim}Ra)$, relative to our *actual* background corpus K_α (which is assumed to contain no

[9] Metaphysically, there may be a problem with "non-natural kinds" (in Quine's sense – *e.g.*, disjunctive and negative properties) participating in certain kinds of causal or other law-like relations. This sort of problem has been suggested in the contemporary literature by Armstrong (1978), Shoemaker (1980), and others. But, we think this metaphysical fact (if it is a fact) has few (if any) *confirmational* consequences. Confirmation is a logical or epistemic relation, which may or may not align neatly with metaphysical relations like causation or law-likeness.

"unkosher" information about instance a). Thus, they aim to show (at least) that relative to some measure c of how strongly evidence supports a hypothesis, the following COMParative claim holds:[10]

$$(\text{COMP}_c)\quad c[(\forall x)(Rx \supset Bx), (Ra \cdot Ba) \mid K_\alpha] > c[(\forall x)(Rx \supset Bx), (\sim Ba \cdot \sim Ra) \mid K_\alpha].$$

Here $c(H, E \mid K)$ is some Bayesian measure of the *degree* to which E confirms H, relative to background corpus K. The typical Bayesian strategy is to isolate constraints on K_α that are as minimal as possible (hopefully, even ones that Hempel would see as kosher), but that guarantee that (COMP_c) obtains.

As it stands, (COMP_c) is somewhat unclear. *Many* different Bayesian relevance measures c have been proposed and defended in the contemporary literature on Bayesian confirmation. The four most popular of these are the following.[11]

---

[10] As Chihara (1981) points out, "there is no such thing as *the* Bayesian solution. There are many different 'solutions' that Bayesians have put forward using Bayesian techniques". That said, we present here what we take to be the most standard assumptions Bayesians tend to make in their handling of the paradox − assumptions that are *sufficient* for the desired comparative and quantitative confirmation-theoretic claims. On this score, we follow Vranas (2004). However, not all Bayesians make precisely these assumptions. To get a sense of the variety of Bayesian approaches, see, *e.g.*: (Alexander 1958); (Chihara 1981); (Earman 1992); (Eells 1982); (Gaifman 1979); (Gibson 1969); (Good 1960, 1961); (Hesse 1974); (Hooker & Stove 1968); (Horwich 1982), (Hosiasson-Lindenbaum 1940); (Howson & Urbach 1993); (Jardine 1965); (Mackie 1963); (Nerlich 1964); (Suppes 1966); (Swinburne 1971, 1973); (Wilson 1964); (Woodward 1985); (Hintikka 1969); (Humburg 1986); (Maher 1999, 2004); (Vranas 2004).

[11] See Fitelson (1999) (2001) for historical surveys. Notice taking *logarithms* of the ratio measures makes them positive in cases of confirmation, negative in cases of disconfirmation, and zero in cases of neutrality of irrelevance. This is a useful convention, but since *log*s don't alter the ordinal structure of the measures, it is a *mere* convention.

- The Difference: $d[H, E \mid K] = \mathrm{P}[H \mid E \cdot K] - \mathrm{P}[H \mid K]$
- The Log-Ratio: $r[H, E \mid K] = \log(\mathrm{P}[H \mid E \cdot K] \,/\, \mathrm{P}[H \mid K])$
- The Log-Likelihood-Ratio: $l[H, E \mid K] = \log(\mathrm{P}[E \mid H \cdot K] \,/\, \mathrm{P}[E \mid {\sim}H \cdot K])$
- The Normalized Difference: $s[H, E \mid K] = \mathrm{P}[H \mid E \cdot K] - \mathrm{P}[H \mid {\sim}E \cdot K]$

Measures *d, r*, and *l* all satisfy the following *desideratum*, for all H, E_1, E_2, and K:

(†) if $\mathrm{P}[H \mid E_1 \cdot K] > \mathrm{P}[H \mid E_2 \cdot K]$, then $c[H, E_1 \mid K] > c[H, E_2 \mid K]$.

But, interestingly, measure *s* does *not* satisfy (†). So, putting *s* aside, if one uses either *d*, *r*, or *l* to measure confirmation, then one can establish the desired comparative claim *simply* by demonstrating that:

($\mathrm{COMP_P}$) $\mathrm{P}[(\forall x)(Rx \supset Bx) \mid Ra \cdot Ba \cdot K_\alpha] > \mathrm{P}[(\forall x)(Rx \supset Bx) \mid {\sim}Ba \cdot {\sim}Ra \cdot K_\alpha]$

(If one uses *s*, then one has a bit more work to do to establish the desired comparative conclusion, because ($\mathrm{COMP_P}$) does not entail (COMP_s).)[12]

Some Bayesians go farther than this by trying to establish not only the *comparative* claim (COMPc), but also the *quantitative* claim that the observation of a non-black non-raven confirms "All ravens are black" to a very minute degree. That is, in addition to the comparative claim, some Bayesians also go for the following QUANTative claim:

(QUANT_c) $c[(\forall x)(Rx \supset Bx), ({\sim}Ba \cdot {\sim}Ra) \mid K_\alpha] > 0$, but very nearly 0.

[12] This has led some former defenders of *s* to abandon it as a measure of incremental confirmation. See Joyce (2004, *fn*. 11). See, also, Eells and Fitelson (2000, 2002) and Fitelson (2001) for further peculiarities of the measure *s*.

Let's begin by discussing the canonical contemporary Bayesian *comparative* analysis of the paradox. In essence, almost all such accounts trade on the following three assumptions about K_α (where we may suppose that the object *a* is sampled at random from the universe):[13]

(1) $P[{\sim}Ba \mid K_\alpha] > P[Ra \mid K_\alpha]$

(2) $P[Ra \mid (\forall x)(Rx \supset Bx) \cdot K_\alpha] = P[Ra \mid K_\alpha]$

(3) $P[{\sim}Ba \mid (\forall x)(Rx \supset Bx) \cdot K_\alpha] = P[{\sim}Ba \mid K_\alpha]$

Basically, assumption (1) relies on our knowledge that (according to K_α) there are more non-black objects in the universe than there are ravens. This seems like a very plausible distributional constraint on K_α, since – as far as we *actually know* – it is true. Assumptions (2) and (3) are more controversial. We will say more about them shortly. First, we note an important and pretty well-known theorem.

THEOREM. (1)–(3) entails (COMP_P). Therefore, since *d*, *r*, and *l* each satisfy (†), it follows that (1)–(3) entails (COMP_d), (COMP_r), and (COMP_l).

---

[13] Often, Bayesians use a *two-stage sampling* model in which *two* objects *a* and *b* are sampled at random from the universe, and where K_α entails $(Ra \cdot {\sim}Bb)$ (*e.g.*, Earman 1992). On that model we still have (2), but (3) is replaced with $P[{\sim}Bb \mid H \cdot K_\alpha] = P[{\sim}Bb \mid K_\alpha]$, and ($\mathrm{COMP}_\mathrm{P}$) is replaced by ($\mathrm{COMP}_\mathrm{P}'$) $P[H \mid Ra \cdot Ba \cdot K_\alpha] > P[H \mid {\sim}Bb \cdot {\sim}Rb \cdot K_\alpha]$. However, no real loss of generality comes from restricting our treatment to "one-stage sampling" – i.e., to the selection of a single object *a*, which K_α doesn't specify to be either an *R* or a *~B* (Vranas 2004, *fns*. 10 and 18). We prefer a *one*-stage sampling approach because it simplifies the analysis somewhat, and because we think it is closer in spirit to what Hempel and Goodman took the original paradox to be about – where K_α is assumed *not* to have any implications about the color or species of the objects sampled, and where a single object is observed "simultaneously" for its color and species.

In fact, (1)–(3) entails *much more* than ($\mathrm{COMP_P}$), as the following theorem illustrates:

THEOREM. (1)–(3) also entail the following:

(4) $\mathrm{P}[(\forall x)(Rx \supset Bx) \mid {\sim}Ba \cdot {\sim}Ra \cdot K_\alpha] > \mathrm{P}[(\forall x)(Rx \supset Bx) \mid K_\alpha]$

(5) $s[(\forall x)(Rx \supset Bx), (Ra \cdot Ba) \mid K_\alpha] > s[(\forall x)(Rx \supset Bx), ({\sim}Ba \cdot {\sim}Ra) \mid K_\alpha]$

In other words, (4) tells us that assumptions (1)–(3) entail that the observation of a non-black non-raven *positively confirms* that all ravens are black – *i.e.*, that the paradoxical conclusion (PC) is true. And, (5) tells us that even according to measure *s* (a measure that *violates* (†)) the observation of a black raven confirms that all ravens are black more strongly than the observation of a non-black non-raven.

The fact that (1)–(3) entail (4) and (5) indicates that the canonical Bayesian assumptions go far beyond the minimal comparative claim most Bayesians were looking for. Why, for instance, should a Bayesian be committed to the *qualitative* paradoxical conclusion (PC)? After all, as Patrick Maher and I.J. Good have made so clear, probabilists don't *have* to be committed to qualitative claims like (NC) and (PC). It would be nice (and perhaps more informative about the workings of Bayesian confirmation) if there were assumptions weaker than (1)–(3) that sufficed to establish (*just*) the *comparative* claim ($\mathrm{COMP_P}$), while implying no commitment to specific *qualitative* claims like (PC). Happily, there are such weaker conditions. But, before we turn to them, we first need to briefly discuss the *quantitative* Bayesian approaches as well.

Various Bayesians go farther than (COMP_c) in their analysis of the ravens paradox. They seek to identify stronger constraints, stronger background knowledge K_α, that entails *both* (COMP_c) *and* (QUANT_c). The most common strategy along these lines is simply to strengthen assumption (1), as follows:

(1′) $P[{\sim}Ba \mid K_\alpha] \gg P[Ra \mid K_\alpha]$ – e.g., because there are *far* fewer ravens than non-black things in the universe.

Peter Vranas (2004) provides a very detailed discussion of quantitative Bayesian approaches to the ravens paradox along these lines. We won't dwell too much on the details of these approaches here. Vranas has already done an excellent job of analyzing them. However, some brief remarks on a result Vranas proves (and uses in his analysis) are worth considering.

Vranas shows that assumptions (1′) and (3) (without (2)) are *sufficient* for (QUANT_c) to hold – i.e. for $(\forall x)(Rx \supset Bx)$ to be positively confirmed by $({\sim}Ba \cdot {\sim}Ra)$ given K_α, but only by a very small amount. He shows this for all four measures of confirmation d, r, l, and s. Moreover, he argues that in the presence of (1′), (3) is "approximately necessary" for (QUANT_c). That is, he *proves* that given (1′), and supposing that $P[H \mid K_\alpha]$ is not too small, the following approximate claim is *necessary* for (QUANT_c):

(3′) $P[{\sim}Ba \mid (\forall x)(Rx \supset Bx) \cdot K_\alpha] \approx P[{\sim}Ba \mid K_\alpha]$.

Vranas then argues that Bayesians have given no good reason for *assuming* this (necessary and sufficient) condition. Thus, he concludes, Bayesian resolutions of the paradox that claim non-black non-ravens confirm by a tiny bit, due to assumption (1'), have failed to establish a condition they must employ to establish this claim – they have failed to establish (3').[14]

[14] However, Vranas does not argue that (3′) is false or implausible – only that no good argument for its plausibility has been given. So, it is consistent with his result that one might be able to find some plausible condition X that, together with (1′), implies (QUANT_c). Vranas' result would then show that condition X (together with (1′)) also implies (3′) – and so in effect would provide a plausibility argument for (3′). Some of the results we prove in the next two sections will provide such conditions, X.

Vranas' claim that (3) is "approximately necessary" for (QUANT$_c$) may be somewhat misleading. It makes it sound as if (3) has a certain property. But, in fact, nothing about (3) itself follows from Vranas' results. It is more accurate to say (as Bob Dylan might) that "approximately (3)" (*i.e.*, (3′)) is *necessary* for (QUANT$_c$). To see the point, note that (3) is a rather strong *independence* assumption, which entails many other identities, including:

(3.1) $P[(\forall x)(Rx \supset Bx) \mid Ba \cdot K_\alpha] = P[(\forall x)(Rx \supset Bx) \mid K_\alpha]$, and

(3.2) $P[(\forall x)(Rx \supset Bx) \mid Ba \cdot K_\alpha] = P[(\forall x)(Rx \supset Bx) \mid {\sim}Ba \cdot K_\alpha]$.

But, (3′) is *not* an independence assumption. Indeed, (3′) is far weaker than an independence assumption, and it does *not* entail the parallel *approximates*:

(3′.1) $P[(\forall x)(Rx \supset Bx) \mid Ba \cdot K_\alpha] \approx P[(\forall x)(Rx \supset Bx) \mid K_\alpha]$, and

(3′.2) $P[(\forall x)(Rx \supset Bx) \mid Ba \cdot K_\alpha] \approx P[(\forall x)(Rx \supset Bx) \mid {\sim}Ba \cdot K_\alpha]$.

Vranas argues convincingly that strong independence assumptions like (3) [and (2)] have not been well motivated by Bayesians who endorse the quantitative approach to the ravens paradox. He rightly claims that this is a lacuna in the canonical quantitative Bayesian analyses of the paradox. But, what he ultimately shows is somewhat weaker than appearances might suggest. In the next two sections we will describe (*pace* Vranas and most other commentators) considerably weaker sets of assumptions for the comparative and the quantitative Bayesian approaches.

8. A New Bayesian Approach to the Paradox

As we have seen, Bayesians typically make two quite strong independence assumptions in order to establish the comparative claim that a black raven confirms more than does a non-black non-raven. In addition they usually suppose that given only actual background knowledge K_α, a non-black instance is more probable than a raven instance. Happily, there is a quite satisfactory analysis of *the ravens* that employs none of these assumptions up front. This solution to the ravens paradox is more general than any other we know of, and it draws on much weaker assumptions. It *solves* the paradox in that it supplies plausible necessary and sufficient conditions for an instance of a black raven to be *more favorable* to 'All ravens are black' than an instance of a non-black non-raven. Our most general result doesn't depend on whether the Nicod Condition (NC) is satisfied, and does not draw on probabilistic independence. Nor does it assume the more plausible claim that (given background knowledge) a non-black instance is more probable than a raven instance (i.e. assumption (1) in the previous section). Indeed, the conditions for this result may be satisfied even if an instance of a black raven *lowers* the degree of confirmation for 'All ravens are black'. In that case it just shows that non-black non-ravens lower the degree of confirmation even more. Thus, this result strips the Bayesian solution to bare bones, decoupling it from any of the usual assumptions. It then permits additional strengthening via the separate introduction of whatever additional suppositions may seem plausible and fitting (e.g. those that lead to positive confirmation).

For the sake of notational simplicity, let 'H' abbreviate 'All ravens are black' – i.e., '$(\forall x)(Rx \supset Bx)$'. Let '$K$' be a statement of whatever background knowledge you may think relevant – e.g. K might imply, among other things, that ravens exist and that non-black things exist, $((\exists x)Rx \cdot (\exists x){\sim}Bx)$. One object, call it '$a$' will be observed for color and to see whether it is a raven. The idea is to assess, in advance of observing it, whether a's turning out to be a black raven, $(Ra{\cdot}Ba)$, would make H more strongly supported than would a's turning out to be a non-black non-raven, $({\sim}Ra{\cdot}{\sim}Ba)$. We want to find plausible conditions for $P[H \mid Ba{\cdot}Ra{\cdot}K] > P[H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K]$ to hold. Equivalently, we want to find plausible conditions for the ratio

$P[Ba\cdot Ra \mid H\cdot K] / P[Ba\cdot Ra \mid {\sim}H\cdot K]$ to exceed the ratio $P[{\sim}Ba\cdot{\sim}Ra \mid H\cdot K] / P[{\sim}Ba\cdot{\sim}Ra \mid {\sim}H\cdot K]$.[15] We will attack the paradox by finding plausible sufficient and necessary conditions for this relationship between likelihood-ratios.[16] Notice that in general this relationship, $P[Ba\cdot Ra \mid H\cdot K] / P[Ba\cdot Ra \mid {\sim}H\cdot K] > P[{\sim}Ba\cdot{\sim}Ra \mid H\cdot \mathrm{K}] / P[{\sim}Ba\cdot{\sim}Ra \mid {\sim}H\cdot K]$, may hold regardless of whether the instance ($Ba\cdot Ra$) raises the confirmation of the hypothesis – i.e., regardless of whether $P[H \mid Ba\cdot Ra\cdot K]$ is greater than, or less than, $P[H \mid K]$.[17] Thus, no condition that implies black ravens *raise* the degree of confirmation can be a *necessary condition* for black ravens to yield greater support than non-black non-ravens. Any such *positive confirmation implying condition* goes beyond what is strictly needed here.

---

[15] $P[H \mid Ba\cdot Ra\cdot K] > P[H \mid {\sim}Ba\cdot{\sim}Ra\cdot K]$ iff $P[{\sim}H \mid Ba\cdot Ra\cdot K] < P[{\sim}H \mid {\sim}Ba\cdot{\sim}Ra\cdot K]$. So, $P[H \mid Ba\cdot Ra\cdot K] > P[H \mid {\sim}Ba\cdot{\sim}Ra\cdot K]$ iff $P[H \mid Ba\cdot Ra\cdot K] / P[{\sim}H \mid Ba\cdot Ra\cdot K] > P[H \mid {\sim}Ba\cdot{\sim}Ra\cdot K] / P[{\sim}H \mid {\sim}Ba\cdot{\sim}Ra\cdot K]$ iff $P[Ba\cdot Ra \mid H\cdot K] / P[Ba\cdot Ra \mid {\sim}H\cdot K] > P[{\sim}Ba\cdot{\sim}Ra \mid H\cdot \mathrm{K}] / P[{\sim}Ba\cdot{\sim}Ra \mid {\sim}H\cdot K]$.

[16] Throughout the remainder of the paper our treatment will focus on the relationship between these likelihood-ratios. However, for $0 < P[H \mid K] < 1$, we have $P[H \mid Ba\cdot Ra\cdot K] > P[H \mid {\sim}Ba\cdot{\sim}Ra\cdot K]$ if and only if $c[H, Ba\cdot Ra \mid K] > c[H, {\sim}Ba\cdot{\sim}Ra \mid K]$, where c is any of the three measures of incremental confirmation d, r, and l. This is the result (†) discussed in the previous section, together with its (easy to established) converse. So, a specific qualitative relationship (>, or =, or <) holds between these likelihood-ratios just in case it holds between $P[H \mid Ba\cdot Ra\cdot K]$ and $P[H \mid {\sim}Ba\cdot{\sim}Ra\cdot K]$, just in case it holds between $c[H, Ba\cdot Ra \mid K]$ and $c[H, {\sim}Ba\cdot{\sim}Ra \mid K]$, where c is any of the measures d, r, and l.

[17] That is, the conditions we will establish do not imply that likelihood-ratio $P[Ba\cdot Ra \mid H\cdot K] / P[Ba\cdot Ra \mid {\sim}H\cdot K]$ is itself greater than 1. And, since this likelihood-ratio will be greater than 1 *just when* H receives positive support from ($Ba\cdot Ra$) (i.e. *just when* $P[H \mid Ba\cdot Ra\cdot K] > P[H \mid K]$), it follows that we will not be requiring that H receive positive support from ($Ba\cdot Ra$). (See Claim 2 in the Appendix for more about this.)

We assume throughout the remainder of the paper the following very weak and highly plausible *non-triviality conditions*:

Non-triviality Assumptions: $P[Ba{\cdot}Ra \mid K] > 0$, $P[{\sim}Ba{\cdot}{\sim}Ra \mid K] > 0$, $P[{\sim}Ba{\cdot}Ra \mid K] > 0$, $0 < P[H \mid Ba{\cdot}Ra{\cdot}K] < 1$, $0 < P[H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K] < 1$.[18]

That is, we assume that it is at least epistemically (confirmationally) possible, given only background *K*, that observed object *a* will turn out to be a black raven; and possible that *a* will turn out to be a non-black non-raven; and even possible that *a* will turn out to be a non-black raven – a falsifying instance of *H*. Furthermore, we assume that finding *a* to be a black raven neither absolutely *proves* nor absolutely *falsifies* 'All ravens are black', nor does finding *a* to be a non-black non-raven do so.

Our analysis of *the ravens* will draw on three factors, which we label 'p', 'q', and 'r'.

Definition: Define $q = P[{\sim}Ba \mid {\sim}H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K]$, define $r = P[{\sim}Ba \mid H{\cdot}K] / P[Ra \mid H{\cdot}K]$, and define $p = P[Ba \mid Ra{\cdot}{\sim}H{\cdot}K]$.

Given *Non-triviality*, p, q, and r are well-defined (q and r have non-zero denominators); q and r are greater than 0; and p is greater than 0 and less than 1. (See Lemma 1 in the Appendix.)

The factor r represents how much more likely it is that *a* will be a non-black thing than be a raven *if* the world in fact contains only black ravens (i.e. if *H* is true). Given the kind of world we think we live in, r should be quite large, since even if all of the ravens are black, the non-black things far outnumber the ravens. Similarly, the factor q represents how much more likely it is that *a* will be a non-black thing than be a raven *if* the world in fact contains non-black ravens (i.e. if *H* is false). Given the kind of world we think we live in, q should also be quite large, since the

[18] $P[Ba{\cdot}Ra \mid K] > 0$ and $P[{\sim}Ba{\cdot}{\sim}Ra \mid K] > 0$ are required for $P[H \mid Ba{\cdot}Ra{\cdot}K]$ and $P[H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K]$ to be well-defined; $0 < P[H \mid Ba{\cdot}Ra{\cdot}K] < 1$ implies $0 < P[H \mid K] < 1$. Other implication of *Non-trivality* are in Appendix Lemma 1.

non-black things far outnumber the ravens even if some of the non-black things happen to be ravens. However, though plausibly r and q are very large, for now we will assume neither this nor anything else about their values except what is implied by the Non-triviality Assumptions – i.e. that r and q are well-defined and greater than 0.

Suppose that *H* is in fact false – i.e. non-black ravens exist – and suppose that *a* is a raven. How likely is it that *a* will turn out to be black? The factor p represents this likelihood. This factor may be thought of as effectively representing a "mixture" of the likelihoods due to the various possible alternative hypotheses about the frequency of black birds among the ravens. It would be reasonable to suppose that the value of p is pretty close to 1 – if there are non-black ravens, their proportion among all ravens is most plausibly some small percentage; so the proportion of black birds among ravens should be a fairly small increment below 1. However, for now we will not assume this, or anything else about the value of p, except what is implied by the Non-triviality Assumptions – i.e. that $0 < p < 1$ (shown in Lemma 1 of the Appendix).

It turns out that the relative confirmational support for *H* from to a black raven instance as compared to that from a non-black non-raven instance is merely a function of p, q, and r.

Theorem 1: Given *Non-triviality*, it follows that $q > (1-p) > 0$ and

$$\frac{P[Ba{\cdot}Ra \mid H{\cdot}K] \,/\, P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]}{P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K] \,/\, P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]} = [q-(1-p)]/(p{\cdot}r) > 0.$$

(This and the other theorems are proved in the Appendix.)

This theorem does not itself express the necessary and sufficient conditions for black ravens to favor 'All ravens are black' more strongly than do non-black non-ravens. But an obvious Corollary does so.

Corollary 1: Given *Non-triviality*,

$$\frac{P[Ba \cdot Ra \mid H \cdot K] \,/\, P[Ba \cdot Ra \mid {\sim}H \cdot K]}{P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] \,/\, P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K]} \;>\; 1 \text{ if and only if } q - (1-p) > p \cdot r.$$

And, more generally, for any real number s,

$$\frac{P[Ba \cdot Ra \mid H \cdot K] \,/\, P[Ba \cdot Ra \mid {\sim}H \cdot K]}{P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] \,/\, P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K]} \;=\; s \;=\; [q-(1-p)]/(p \cdot r) > 1 \text{ if and only if}$$

$$[q - (1-p)] \;=\; s \cdot p \cdot r \;>\; p \cdot r.$$

This gives us a fairly useful handle on what it takes for a black raven to support *H* more than a non-black non-raven. For instance, suppose that q = r. Then the corollary implies that the value of the ratio of likelihood-ratios is greater than 1 *just in case* q = r > 1.[19] Thus, if the likelihood that an object is non-black is greater than the likelihood that it's a raven, and is greater by the same amount regardless of whether or not every raven is black, then a black raven supports 'All ravens are black' more strongly than does a non-black non-raven.

Notice that none of this depends on either *Ba* or *Ra* being probabilistically independent of *H*. Such independence, if it held, would make $P[{\sim}Ba \mid {\sim}H \cdot K] = P[{\sim}Ba \mid H \cdot K] = P[{\sim}Ba \mid K]$ and make $P[Ra \mid {\sim}H \cdot K] = P[Ra \mid H \cdot K] = P[Ra \mid K]$. In that case we would indeed have q = r, and so the result discussed in the previous paragraph would apply. However, that result applies even in cases where probabilistic independence fails miserably – even when $P[{\sim}Ba \mid {\sim}H \cdot K]/P[{\sim}Ba \mid H \cdot K]$ is very far from 1, provided only that $P[Ra \mid {\sim}H \cdot K]/P[Ra \mid H \cdot K]$ is equally far from 1.

What if q ≠ r? Theorem 1 tell us that q > (1−p) > 0, so q−(1−p) is positive and a little smaller than q itself. As long as this q−(1−p) remains larger than r, the corollary tells us that the likelihood-ratio due to a black raven favors *H* more than does the likelihood-ratio due to a non-

[19] Since for q = r, s > 1 iff q−(1−p) > p·q iff q·(1−p) = q − p·q > (1−p) iff q > 1.

black non-raven. Indeed $q-(1-p)$ need only remain larger than a fraction p of r in order to yield the desired result.

It turns out that 1/p is a convenient benchmark for comparing the size of the black-raven likelihood-ratio to the size non-black-non-raven likelihood-ratio.

Corollary 2: Given *Non-triviality*, for real number s such that

$$\frac{P[Ba \cdot Ra \mid H \cdot K] \,/\, P[Ba \cdot Ra \mid {\sim}H \cdot K]}{P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] \,/\, P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K]} = s = [q-(1-p)]/(p \cdot r), \text{ we have the following:}$$

(1) $s > (1/p) > 1$ iff $q-(1-p) > r$

(2) $s = (1/p) > 1$ iff $q-(1-p) = r$

(3) $(1/p) > s > 1$ iff $r > q-(1-p) > p \cdot r$.

Notice that when $q = r$, Clause 3 applies (because then $r > q-(1-p)$); so the value of the ratio of the likelihood-ratios, s, must be strictly between 1/p and 1. Alternatively, when q diminished by $(1-p)$ is greater than r, Clause 1 applies; so the ratio of likelihood-ratios s must be greater than $(1/p)$, possibly much greater. Indeed, looking back at Corollary 1, we see that the value of the ratio of likelihood ratios s can be enormous, provided only that $[q-(1-p)] >>> (p \cdot r)$.

The emergence of 1/p as a particularly useful benchmark is no accident. For, p is just $P[Ba \mid Ra \cdot {\sim}H \cdot K]$, so $1/p = P[Ba \mid Ra \cdot H \cdot K] / P[Ba \mid Ra \cdot {\sim}H \cdot K]$. Furthermore, if the usual independence assumption (2) were to hold (i.e. if $P[Ra \mid H \cdot K] = P[Ra \mid K]$), it would follow that $P[Ra \mid H \cdot K] = P[Ra \mid {\sim}H \cdot K]$; and then we'd have $1/p = P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid {\sim}H \cdot K]$. Following this thought further, the usual Bayesian analysis adds independence assumption (3) (i.e. $P[{\sim}Ba \mid H \cdot K] = P[{\sim}Ba \mid K]$) to get $P[{\sim}Ba \mid H \cdot K] = P[{\sim}Ba \mid {\sim}H \cdot K]$; from which we'd have $P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] / P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K] = P[{\sim}Ba \mid H \cdot K] / (P[{\sim}Ra \mid {\sim}Ba \cdot {\sim}H \cdot K] \cdot P[{\sim}Ba \mid {\sim}H \cdot K]) = 1/P[{\sim}Ra \mid {\sim}Ba \cdot {\sim}H \cdot K]$, where $P[Ra \mid {\sim}Ba \cdot {\sim}H \cdot K]$ should be just a smidgen, ε, above 0 – because, very probably, only a really minuscule proportion of the non-black things are ravens, regardless of whether H is true or false. Thus, the usual analysis would peg the *ratio of likelihood-ratios* at a value $s = (1-\varepsilon)/p$ (for ε almost 0), which is just a tiny bit below $1/p$ – which is only within the range of possible values for s encompassed by Clause 3 of Corollary 2, and merely within the uppermost end of that range. In light of this, the benchmark $1/p$ in Corollary 2 provides a telling indicator of the extent to which our treatment supersedes the usual approach.

Theorem 1 and its Corollaries show that for a very wide range of probabilistic confirmation functions P, a black raven is *more confirming* of 'All ravens are black' than is a non-black non-raven. These functions are so diverse that some of them even permit a black raven to provide evidence against 'All ravens are black' (i.e. make $P[Ba{\cdot}Ra \mid H{\cdot}K] / P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K] < 1$). Only a small range of these functions abide by the usual independence claims. For black ravens to be *more confirming*, all that matters are the relative sizes of q and r, as mediated by the factor p.

Let's now look at one more theorem that *solves* the paradox by drawing on additional conditions that restrict the values of q and r in a plausible way. This result is less general than Theorem 1 and its corollaries, but closely related to them.[20]

Theorem 2: Given *Non-triviality*, both of the following clauses hold:

(2.1) If $P[{\sim}Ba \mid H{\cdot}K] > P[Ra \mid H{\cdot}K]$ (i.e. if $r > 1$) and

$O[H \mid Ra{\cdot}K] / O[H \mid {\sim}Ba{\cdot}K] > p + (1-p)/r$, then

$$\frac{P[Ba{\cdot}Ra \mid H{\cdot}K] / P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]}{P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K] / P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]} > 1.$$

(2.2) If $P[{\sim}Ba \mid H{\cdot}K] \leq P[Ra \mid H{\cdot}K]$ (i.e. $r \leq 1$), but either $P[{\sim}Ba \mid K] > P[Ra \mid K]$

[20] One clause of this result draws on the notion of *odds*, O. By definition, $O[X \mid Y] = P[X \mid Y] / P[{\sim}X \mid Y]$.

or (at least) $P[\sim Ba \mid \sim H{\cdot}K] > P[Ra \mid \sim H{\cdot}K]$ (i.e. $q > 1$), then

$$\frac{P[Ba{\cdot}Ra \mid H{\cdot}K] \,/\, P[Ba{\cdot}Ra \mid \sim H{\cdot}K]}{P[\sim Ba{\cdot}\sim Ra \mid H{\cdot}K] \,/\, P[\sim Ba{\cdot}\sim Ra \mid \sim H{\cdot}K]} > 1.$$

Clause (2.1) is the more interesting case, since its antecedent conditions are a better fit to the way we typically judge our world to be. The first antecedent of (2.1) draws on the idea that, provided all of the ravens are black, a randomly selected object *a* is more likely (in our world) to be a non-black thing than a raven. This seems really quite plausible. Indeed, not only does it seem that $P[\sim Ba \mid H{\cdot}K]$ is *merely greater than* $P[Ra \mid H{\cdot}K]$, quite plausibly $P[Ra \mid H{\cdot}K]$ is close enough to 0 that $P[\sim Ba \mid H{\cdot}K]$ is billions of times greater than $P[Ra \mid H{\cdot}K]$ (though the theorem itself doesn't suppose *that*).

Now consider the second antecedent to (2.1). One wouldn't normally think that the mere fact that an object is black (without also taking account of whether it's a raven) should provide *more evidence* for 'All ravens are black' than would the mere fact that an object is a raven (without taking account of its color). Indeed, generally speaking, one would expect $O[H \mid Ra{\cdot}K]$ to be very nearly equal to $O[H \mid \sim Ba{\cdot}K]$. However, the second condition for Clause (2.1) is even weaker than this. Notice that for $r > 1$ the term $p + (1-p)/r$ is less than $p+(1-p) = 1$; and the larger r happens to be (i.e. the greater the ratio $r = P[\sim Ba \mid H{\cdot}K] \,/\, P[Ra \mid H{\cdot}K]$ is), the smaller $p + (1-p)/r$ will be, approaching the lower bound $p = P[Ba \mid Ra{\cdot}\sim H{\cdot}K]$ for very large r. Thus, the second condition for (2.1) will be satisfied provided that *either* $O[H \mid Ra{\cdot}K]$ is bigger than or equal to $O[H \mid \sim Ba{\cdot}K]$ (perhaps much bigger) *or* $O[H \mid Ra{\cdot}K]$ is a bit smaller than $O[H \mid \sim Ba{\cdot}K]$. Thus, this second condition can *fail to hold* only if (without taking account of whether it's a raven) a black object provides *more than a bit more evidence* for 'All ravens are black' than would a raven (without taking account of its color).[21]

Although the antecedent conditions for Clause (2.2) seem a less plausible fit to our world, it fills out Theorem 2 in an interesting way. Think of it like this. It is reasonable to suppose, given plausible background knowledge, that the non-black things will be much more numerous than ravens, regardless of whether all the ravens are black. But perhaps this intuition is confused. It is clearly guided by the fact that we inhabit a world in which there are far more non-black things than ravens. Problem is, if our world is one in which there are non-black ravens, we may only be warranted in taking the non-black things to outnumber the ravens in worlds like our – i.e. worlds where H is false. If, on the other hand, ours happens to be a world in which all of the ravens are black, then we may only be warranted in taking the non-black things to outnumber the ravens in worlds like ours – i.e. worlds where H is true. But we don't know which of these two kinds of worlds ours happens to be. That is precisely what is at issue – precisely what the evidence is supposed to tell us. Nevertheless, we can easily fineness this apparent difficulty. For, the apparent dilemma takes it as granted that *either* non-black things are much more numerous than

---

[21] An equivalent (and perhaps more illuminating) alternative to the second condition for Clause (2.1) is this: the ratio $P[Ra \mid H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K]$ is no less than the ratio $P[{\sim}Ba \mid H{\cdot}K] / P[{\sim}Ba \mid {\sim}H{\cdot}K]$, or perhaps only a bit less – i.e. $(P[Ra \mid H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K]) / (P[{\sim}Ba \mid H{\cdot}K] / P[{\sim}Ba \mid {\sim}H{\cdot}K]) \geq (p+(1-p)/r)$. Here $p+(1-p)/r < 1$ because the first condition of Clause 2.1 requires $r > 1$. This condition (and the equivalent odds condition) is *strictly weaker* than the usual independence assumptions. For, if independence assumption (2) holds, then the $P[Ra \mid H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K] = 1$, and if independence assumption (3) holds, then the $P[{\sim}Ba \mid H{\cdot}K] / P[{\sim}Ba \mid {\sim}H{\cdot}K] = 1$. Thus, the two usual conditions entail the much more restrictive $P[Ra \mid H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K] = P[{\sim}Ba \mid H{\cdot}K] / P[{\sim}Ba \mid {\sim}H{\cdot}K]$ – i.e. $O[H \mid Ra{\cdot}K] = O[H \mid {\sim}Ba{\cdot}K]$.

ravens if H holds, *or* non-black things are much more numerous than ravens if $\sim H$ holds. Thus, given reasonable background knowledge, for an object a about which nothing else is known, *either* $P[\sim Ba \mid H{\cdot}K] > P[Ra \mid H{\cdot}K]$ (i.e. $r > 1$) *or* $P[\sim Ba \mid \sim H{\cdot}K] > P[Ra \mid \sim H{\cdot}K]$ (i.e. $q > 1$) (or perhaps $P[\sim Ba \mid K] > P[Ra \mid K]$). But Clause (2.1) of the theorem already takes account of the case where $P[\sim Ba \mid H{\cdot}K] > P[Ra \mid H{\cdot}K]$. So Clause (2.2) deals with the remaining case: that in case $P[\sim Ba \mid H{\cdot}K] \leq P[Ra \mid H{\cdot}K]$ (i.e. $r \leq 1$) holds, at least $P[\sim Ba \mid \sim H{\cdot}K] > P[Ra \mid \sim H{\cdot}K]$, or maybe $P[\sim Ba \mid K] > P[Ra \mid K]$ holds. This is the only condition required for (2.2), and it's a very weak condition indeed.

Consider the disjunction of the antecedent conditions for Clause (2.1) with the antecedent conditions for Clause (2.2). This disjunction is a highly plausible claim – even more plausible than each antecedent taken alone. Given realistic background knowledge K, any reasonable probabilistic confirmation function P should surely satisfy the full antecedent of at least one of these two clauses. Thus, a black raven should *favor* 'All ravens are black' *more* than a non-black non-raven over a *very wide range* of circumstances. Furthermore, neither of the usual *approximate independence conditions* is required for this result. Thus, Theorem 1 and its corollaries together with Theorem 2 dissolve any air of a *qualitative paradox* in the case of *the ravens*.

9. Quantitative Results

Traditional *quantitative* Bayesian approaches also make rather strong independence-like assumptions. For example, in order to establish that a non-black non-raven positively confirms 'All ravens are black' by (only) a very small amount – the thesis we've labeled (QUANT_c), $c[H, \sim Ba{\cdot}\sim Ra \mid K] > 0$ but very near 0 – the usual approach employs an (at least approximate) independence assumption like (3) or (3′), $P[{\sim}Ba \mid H{\cdot}K] \approx P[{\sim}Ba \mid K]$, together with an assumption like (1′), $P[{\sim}Ba \mid K_\alpha] \gg P[Ra \mid K]$.[22]

Quantitative claims like (QUANT_c) are most informative when cashed out in terms of a specific measure of confirmation *c*. That is, although several of the well-studied measures of incremental confirmation (*d*, *r*, and *l*) agree with regard to *qualitative* confirmational relationships, their *quantitative* scales differ in ways that that make quantitative results difficult to compare meaningfully across measures. So in this section we'll restrict our discussion to a single measure of incremental confirmation. In our judgment the most suitable Bayesian measure of incremental confirmation is the (log) likelihood-ratio measure.[23] We have detailed reasons for this assessment (see Fitelson 2001 and 2004), but we'll not pause to discuss them here. Let's see what the likelihood-ratio measure can tell us *quantitatively* about *the ravens*.

In terms of the likelihood-ratio measure, and drawing on our factors p, q, and r, a reworking of Vranas's (2004) result leads to the following:

Theorem 3: If the degree to which a non-black non-raven incrementally confirms 'All ravens are black', as measured by the likelihood-ratio, is in the interval $1 < P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K] / P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K] \leq 1+\varepsilon$, for very small $\varepsilon > 0$, then

$([q-(1-p)]/q) < P[{\sim}Ba \mid H{\cdot}K] / P[{\sim}Ba \mid {\sim}H{\cdot}K] \leq ([q-(1-p)]/q){\cdot}(1+\varepsilon)$.

If instead $(1-\varepsilon) < P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K] / P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K] \leq 1$, then

$([q-(1-p)]/q){\cdot}(1-\varepsilon) < (P[{\sim}Ba \mid H{\cdot}K] / P[{\sim}Ba \mid {\sim}H{\cdot}K]) \leq ([q-(1-p)]/q)$.

---

[22] Vranas (2004) provides a detailed exposition.

[23] We'll suppress the "log", since nothing we'll say depends on the re-scaling of likelihood-ratios by taking the log.

In both cases, for large q, $([q-(1-p)]/q) \approx 1$, so $P[\sim Ba \mid H \cdot K] / P[\sim Ba \mid \sim H \cdot K] \approx 1$.[24]

(Recall that $q = P[\sim Ba \mid \sim H \cdot K] / P[Ra \mid \sim H \cdot K]$, which is plausibly quite large.)

So, the approximate independence of *Ba* from the truth or falsehood of *H*, given *K*, is a necessary condition for a non-black non-raven to provide only a very small amount of positive (or negative) support for 'All ravens are black'. Vranas's point is that traditional Bayesian treatments of the ravens paradox almost always employ the "small positive confirmation from non-black non-ravens" idea, and they inevitably draw *directly* on some such independence assumption to achieve it. But, Vranas argues, no plausible justification for *assuming* this (near) independence has yet been given by those who employ it.

Our approach sidesteps this issue completely. None of our results have relied on assuming approximate independence; indeed, our results haven't even supposed that non-black non-ravens should yield positive confirmation for *H*, either small or large. We've only given sufficient (and necessary) conditions for a black raven to *confirm H more* than would a non-black non-raven.

In order to address the ravens in quantitative terms, let's consider the sizes of $r = P[\sim Ba \mid H \cdot K] / P[Ra \mid H \cdot K]$ and of $q = P[\sim Ba \mid \sim H \cdot K] / P[Ra \mid \sim H \cdot K]$. Given background *K* that reflects how all of us generally believe our world to be, both r and q should presumably be quite large, and should be very nearly the same size. However, notice that such suppositions about r

[24] This approximate independence condition implies approximate independence condition (1'), since $P[\sim Ba \mid K] = P[\sim Ba \mid H \cdot K] \cdot P[H \mid K] + P[\sim Ba \mid \sim H \cdot K] \cdot (1 - P[H \mid K]) \approx P[\sim Ba \mid H \cdot K] \cdot P[H \mid K] + P[\sim Ba \mid H \cdot K] \cdot (1 - P[H \mid K]) = P[\sim Ba \mid H \cdot K]$. The two versions of approximate independence are equivalent if $P[H \mid K]$ isn't extremely close to 1.

and q, even if we take q to precisely equal r, don't imply the approximately independence of either *Ra* or of *Ba* from *H* or from ~*H* (given *K*).[25]

Under such circumstances, let's consider *how much more* a black raven confirms 'All ravens are black' than does a non-black non-raven.

Theorem 4: Given *Non-triviality*, suppose $P[{\sim}Ba \mid H{\cdot}K] / P[Ra \mid H{\cdot}K] \geq L > 1$ (i.e. $r \geq L > 1$); and suppose $P[{\sim}Ba \mid {\sim}H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K]$ is very nearly the same size as r – i.e., for some $\delta > 0$ but near 0, $0 < 1-\delta \leq (P[{\sim}Ba \mid {\sim}H{\cdot}K]/P[Ra \mid {\sim}H{\cdot}K]) / (P[{\sim}Ba \mid H{\cdot}K]/P[Ra \mid H{\cdot}K]) \leq 1+\delta$ (that is, $1-\delta \leq q/r \leq 1+\delta$). Then the "ratio of likelihood ratios" is bounded as follows:

$$(4.1) \quad [(1-\delta) - (1-p)/L]\cdot(1/p) < \frac{P[Ba{\cdot}Ra \mid H{\cdot}K]/P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]}{P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K]/P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]} < (1+\delta)\cdot(1/p).$$

If in addition $P[Ba \mid Ra{\cdot}{\sim}H{\cdot}K] > 1/2$, then we get an improved lower bound:

$$(4.2) \quad (1-\delta)\cdot(1/p) - 1/L < \frac{P[Ba{\cdot}Ra \mid H{\cdot}K]/P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]}{P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K]/P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]} < (1+\delta)\cdot(1/p)$$

In either case, for large very $L > 1$ and positive δ near 0 the "ratio of likelihood ratios" is almost exactly equal to $(1/p)$.[26]

---

[25] To see this clearly, supposes that $P[Ra \mid H{\cdot}K]$ is larger than $P[Ra \mid {\sim}H{\cdot}K]$ by a very large factor $f > 1$ – i.e. $P[Ra \mid H{\cdot}K] = f \cdot P[Ra \mid {\sim}H{\cdot}K]$ – and suppose that $P[{\sim}Ba \mid H{\cdot}K]$ is larger than $P[{\sim}Ba \mid {\sim}H{\cdot}K]$ by the same factor – i.e. $P[{\sim}Ba \mid H{\cdot}K] = f \cdot P[{\sim}Ba \mid {\sim}H{\cdot}K]$. Then we'd have $r = P[{\sim}Ba \mid H{\cdot}K] / P[Ra \mid H{\cdot}K] = P[{\sim}Ba \mid {\sim}H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K] = q$ even though neither *Ra* nor *Ba* would be anywhere close to independence of *H* or ~*H*. The same goes for $P[Ra \mid {\sim}H{\cdot}K]$ larger than $P[Ra \mid H{\cdot}K]$ and $P[{\sim}Ba \mid {\sim}H{\cdot}K]$ larger than $P[{\sim}Ba \mid H{\cdot}K]$, both by very large factor $f > 1$.

[26] Proof: $(P[Ba{\cdot}Ra \mid H{\cdot}K]/P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]) / (P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K]/P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]) = (q/r)\cdot(1/p) - [(1-p)/p]/r$, by Theorem 1. We get the upper bounds as follow: $(q/r)\cdot(1/p) - [(1-p)/p]/r < (q/r)\cdot(1/p) \leq (1+\delta)\cdot(1/p)$. To get the lower

The larger r is, and the closer the size of q is to the size of r (i.e. the smaller δ is), the closer will be the "ratio of likelihood ratios" to 1/p. And, if instead of being nearly the same size as r, q is significantly larger than r, then q/r is significantly larger than 1 and (according to Theorem 1) the "ratio of likelihood ratios" must nearly be $(q/r)\cdot(1/p)$ (precisely $([q-(1-p)]/r)\cdot(1/p)$), which must be significantly larger than 1/p.

Let's illustrate this theorem by plug in some fairly realistic numbers. Suppose, as seems plausible, that r is at least as large as $L = 10^9$. (L should really be much larger than this since, given *H·K*, it seems highly probable that there will be trillions of times more non-black things than ravens, not just billions of times more). And suppose that q is very nearly the same size as r – say, within a million of r, $q = r \pm 10^6$, so that $q/r = 1\pm10^{-3}$. Then Theorem 4 tells us that for $P[Ba \mid Ra\cdot{\sim}H\cdot K] = p > 1/2$, the "ratio of likelihood-ratios" is bounded below by $(1-10^{-3})\cdot(1/p) - 1/10^9 = (.999)\cdot(1/p) - 10^{-9}$; and the upper bound is $(1+10^{-3})\cdot(1/p) = (1.001)\cdot(1/p)$. Thus, to three significant figures the "ratio of likelihood ratios is $(1/p) \pm (.001)/p$.

Suppose $P[Ba \mid Ra\cdot{\sim}H\cdot K] = p$ is somewhere around .9 or .95; so (1/p) is somewhere around $1/.9 \approx 1.11$ or $1/.95 \approx 1.05$. Then a single instance of a black raven may not seem to yield a whole lot more support for *H* than a single instance of a non-black non-raven. However, under plausible conditions a sequence of n instances (i.e. of n black ravens, as compared to n non-black non-ravens) will yield a "ratio of likelihood-ratios" on the order of $(1/p)^n$, which blows up significantly for large n. E.g., for n = 100 instances, $(1/.95)^{100} \approx 169$, and $(1/.9)^{100} \approx 37649$ – that

---

bound in (4.1): $(q/r)\cdot(1/p) - [(1-p)/p]/r > (1-\delta)/p - 1/pr \geq [(1-\delta) - 1/L]\cdot(1/p)$. To get the lower bound in (4.2), first notice that for p > 1/2, $[(1-p)/p] < 1$, so $(q/r)\cdot(1/p) - [(1-p)/p]/r \geq (1-\delta)\cdot(1/p) - 1/r > (1-\delta)\cdot(1/p) - 1/L$.

is, for p = .95, 100 black raven instances would yield a likelihood-ratio 169 times higher than would 100 instances of non-black non-ravens.

Nothing in the previous paragraphs draws on the assumption that a non-black non-raven yields (at most) a tiny amount of support for H – i.e. that $P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] / P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K] = (1 \pm \varepsilon)$. But this may be a plausible enough *additional supposition*. When it holds we have the following result.

Theorem 5: Suppose *Non-triviality*, and suppose that r is large and q is very nearly the same size as r in the sense that $(1-\delta) \leq q/r \leq (1+\delta)$, for very small δ (i.e., suppose the conditions for Theorem 4 hold). And suppose, in addition, that the support for H by a non-black non-raven is very small – i.e. $1-\varepsilon \leq P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] / P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K] \leq 1+\varepsilon$ for very small ε. Then the support for H by a black raven must be

$P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid {\sim}H \cdot K] = (1 \pm \delta) \cdot (1 \pm \varepsilon) \cdot (P[Ba \mid Ra \cdot H \cdot K] / P[Ba \mid Ra \cdot {\sim}H \cdot K]) \approx 1/p$,

where, of course, $P[Ba \mid Ra \cdot H \cdot K] / P[Ba \mid Ra \cdot {\sim}H \cdot K] = 1/p$.[27]

Notice that the suppositions of this theorem permit a non-black non-raven to provide absolutely no support for H ($\varepsilon = 0$), or a tiny bit of positive support ($\varepsilon > 0$), or to even provide a tiny bit of evidence against ($\varepsilon < 0$). Here, rather than *assuming* the near probabilistic independence of Ra and Ba from H and $\sim H$ (given K), we've effectively gotten it for free (via Theorem 3), as a consequence of the more plausible direct supposition that non-black non-ravens

[27] Proof: From Theorem 3 we already have that $P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] / P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K] = (1 \pm \varepsilon)$ implies $P[{\sim}Ba \mid H \cdot K] / P[{\sim}Ba \mid {\sim}H \cdot K] = (1 \pm \varepsilon)$. Then $P[Ra \mid H \cdot K] / P[Ra \mid {\sim}H \cdot K] = [(P[Ra \mid H \cdot K] / P[Ra \mid {\sim}H \cdot K]) / (P[{\sim}Ba \mid H \cdot K] / P[{\sim}Ba \mid {\sim}H \cdot K])] \cdot (P[{\sim}Ba \mid H \cdot K] / P[{\sim}Ba \mid {\sim}H \cdot K]) = (q/r) \cdot (1 \pm \varepsilon) = (1 \pm \delta) \cdot (1 \pm \varepsilon)$. And $P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid {\sim}H \cdot K] = (P[Ba \mid Ra \cdot H \cdot K] / P[Ba \mid Ra \cdot {\sim}H \cdot K]) \cdot (P[Ra \mid H \cdot K] / P[Ra \mid {\sim}H \cdot K])$.

don’t confirm much, if at all. This shows how the effect of *near independence* is accommodated by our analysis, if it happens to be implied by some *additional* plausible supposition – e.g. the assessment that no more than a minute amount of confirmation could come from an observation of a single non-black non-raven instance.

Thus, under quite weak, but highly plausible suppositions, a black raven *favors* ‘All ravens are black’ *more* than would a non-black non-raven by about (1/p) – i.e., by about the amount that a black object supports ‘All ravens are black’, given that it is a raven, since

$$P[Ba \mid Ra\cdot H\cdot K] \,/\, P[Ba \mid Ra\cdot{\sim}H\cdot K] = 1/p.$$ [28]

This quantitative result, together with the results derived in Section 8, shows that a careful Bayesian analysis puts the *paradox* of the ravens to rest.

---

[28] The factor $p = P[Ba \mid Ra\cdot{\sim}H\cdot K]$ is a reflection of both likelihoods and priors probabilities for the whole range of alternative hypotheses H_f, where each says that the frequency of black things among ravens, F[Bx, Rx] = f, is a specific fraction f. When p is pretty close to 1, the only alternative hypotheses H_f that can have non-miniscule prior probabilities are those for which f is pretty close to 1 as well. So a single black raven doesn’t provide very much confirmation for H (i.e., only about 1/p, which isn’t much), because it takes a lot of instances to distinguish between H and the alternatives that have f near 1. To see this formally, consider: for each $k \geq 1$ such that $1/(1-p) > k$,

$p = P[Ba \mid Ra\cdot{\sim}H\cdot K] = \sum_{1>f\geq 0} P[Ba \mid Ra\cdot H_f\cdot K]\, P[H_f \mid Ra\cdot{\sim}H\cdot K] = \sum_{1>f\geq 0} f \cdot P[H_f \mid Ra\cdot{\sim}H\cdot K] <$

$\sum_{1>f\geq 1-k\cdot(1-p)} P[H_f \mid Ra\cdot{\sim}H\cdot K] + [1-k\cdot(1-p)]\cdot\sum_{1-k\cdot(1-p)>f\geq 0} P[H_f \mid Ra\cdot{\sim}H\cdot K] =$

$1 - k\cdot(1-p)\cdot\sum_{1-k\cdot(1-p)>f\geq 0} P[H_f \mid Ra\cdot{\sim}H\cdot K]$, so $\sum_{1-k\cdot(1-p)>f\geq 0} P[H_f \mid Ra\cdot{\sim}H\cdot K] < 1/k$; thus $\sum_{1>f\geq 1-k\cdot(1-p)} P[H_f \mid Ra\cdot{\sim}H\cdot K] > (k-1)/k$. For example, for p = .98 and k = 5 we have that $\sum_{1\geq f>.90} P[H_f \mid Ra\cdot{\sim}H\cdot K] > .80$. Indeed, in this case one only gets $\sum_{1\geq f>.90} P[H_f \mid Ra\cdot K]$ to be as small as .80 by making $P[H_{.9} \mid Ra\cdot{\sim}H\cdot K] = .20$ and $P[H_{.999} \mid Ra\cdot{\sim}H\cdot K] = .80$, and $P[H_f \mid Ra\cdot K] = 0$ for all other values of f. If non-zero priors are more evenly distributed throughout the interval between .9 and 1, then $\sum_{1\geq f>.90} P[H_f \mid Ra\cdot{\sim}H\cdot K]$ has to be quite a bit larger than .80 (in order to permit p = .98).

Appendix: Proofs of Various Results.

Claim 1: Given the Non-trivality Assumptions,

$$\frac{P[Ba{\cdot}Ra \mid H{\cdot}K] \,/\, P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]}{P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K] \,/\, P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]} > 1$$

just in case $P[H \mid Ba{\cdot}Ra{\cdot}K] > P[H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K]$.

Proof: Assuming *Non-triviality* we have: $P[H \mid Ba{\cdot}Ra{\cdot}K] > P[H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K]$ *iff* both

$P[H \mid Ba{\cdot}Ra{\cdot}K] > P[H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K]$ and $P[{\sim}H \mid Ba{\cdot}Ra{\cdot}K] < P[{\sim}H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K]$ *iff*

$P[H \mid Ba{\cdot}Ra{\cdot}K]/P[{\sim}H \mid Ba{\cdot}Ra{\cdot}K] > P[H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K]/P[{\sim}H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K]$ *iff*

$(P[Ba{\cdot}Ra \mid H{\cdot}K]/P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]) \cdot (P[H \mid K]/P[{\sim}H \mid K]) >$

$(P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K]/P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]) \cdot (P[H \mid K]/P[{\sim}H \mid K])$ *iff*

$(P[Ba{\cdot}Ra \mid H{\cdot}K]/P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]) \,/\, (P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K]/P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]) > 1.$

□

The following lemma establishes that all of the terms used to define p, q, and r are non-zero.

Lemma 1: Given *Non-triviality*, it follows that $P[Ra \mid H{\cdot}K] > 0$, $P[{\sim}Ba \mid H{\cdot}K] > 0$, $1 > P[Ra \mid {\sim}H{\cdot}K] > 0$, $1 > P[{\sim}Ba \mid {\sim}H{\cdot}K] > 0$, and $1 > P[Ba \mid Ra{\cdot}{\sim}H{\cdot}K] > 0$.

Proof: From *Non-triviality* we have:

(i) $0 < P[H \mid Ba{\cdot}Ra{\cdot}K] = P[Ba{\cdot}Ra \mid H{\cdot}K] \cdot P[H \mid K] \,/\, P[Ba{\cdot}Ra \mid K]$, so $P[Ra \mid H{\cdot}K] = P[Ba{\cdot}Ra \mid H{\cdot}K] > 0$;

(ii) $0 < P[H \mid {\sim}Ba{\cdot}{\sim}Ra{\cdot}K] = P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K] \cdot P[H \mid K] \,/\, P[{\sim}Ba{\cdot}{\sim}Ra \mid K]$, so $P[{\sim}Ba \mid H{\cdot}K] = P[{\sim}Ba{\cdot}{\sim}Ra \mid H{\cdot}K] > 0$;

(iii) $0 < P[{\sim}H \mid Ba{\cdot}Ra{\cdot}K] = P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K] \cdot P[{\sim}H \mid K] \,/\, P[Ba{\cdot}Ra \mid K]$, so $P[Ba \mid {\sim}H{\cdot}K] \geq P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K] > 0$ and $P[Ra \mid {\sim}H{\cdot}K] \geq P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K] > 0$ and $P[Ba \mid Ra{\cdot}{\sim}H{\cdot}K] > 0$;

(iv) $0 < P[{\sim}H \mid {\sim}Ba \cdot {\sim}Ra \cdot K] = P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K] \cdot P[{\sim}H \mid K] / P[{\sim}Ba \cdot {\sim}Ra \mid K]$, so

$P[{\sim}Ba \mid {\sim}H \cdot K] \geq P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K] > 0$ and $P[{\sim}Ra \mid {\sim}H \cdot K] \geq P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K] > 0$.

(v) $0 < P[{\sim}Ba \cdot Ra \mid K] = P[{\sim}Ba \cdot Ra \mid H \cdot K] \cdot P[H \mid K] + P[{\sim}Ba \cdot Ra \mid {\sim}H \cdot K] \cdot P[{\sim}H \mid K] =$

$P[{\sim}Ba \cdot Ra \mid {\sim}H \cdot K] \cdot P[{\sim}H \mid K] < P[{\sim}Ba \cdot Ra \mid {\sim}H \cdot K] \leq P[Ra \mid {\sim}H \cdot K]$, so $0 < P[{\sim}Ba \mid Ra \cdot {\sim}H \cdot K]$,

so $P[Ba \mid Ra \cdot {\sim}H \cdot K] < 1$.

□

The next claim shows how *positive support* for *H* depends on p and q. Our solution of *the ravens* will not depend on *H* receiving *positive support* (as can be seen by comparing this claim to the main Theorem, which will come next). But it's useful and interesting to see what *positive support* requires.

Claim 2: $P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid {\sim}H \cdot K] > 1$ (i.e. *H* is positively supported by $(Ba \cdot Ra)$) if and only if $P[Ra \mid H \cdot K]/P[Ra \mid {\sim}H \cdot K] > p$ (where $p = P[Ba \mid Ra \cdot {\sim}H \cdot K]$); and $P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] / P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K] > 1$ (i.e. *H* is positively supported by $({\sim}Ba \cdot {\sim}Ra)$) if and only if $P[{\sim}Ba \mid H \cdot K]/P[{\sim}Ba \mid {\sim}H \cdot K] > [q - (1-p)]/q$.

Proof: $P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid {\sim}H \cdot K] = P[Ra \mid H \cdot K] / (P[Ba \mid Ra \cdot {\sim}H \cdot K] \cdot P[Ra \mid {\sim}H \cdot K])$

$= (1/p) \cdot P[Ra \mid H \cdot K]/P[Ra \mid {\sim}H \cdot K]$

> 1 iff $P[Ra \mid H \cdot K]/P[Ra \mid {\sim}H \cdot K] > p$.

$P[{\sim}Ba \cdot {\sim}Ra \mid H \cdot K] / P[{\sim}Ba \cdot {\sim}Ra \mid {\sim}H \cdot K]$

$= P[{\sim}Ba \mid H \cdot K] / (P[{\sim}Ba \mid {\sim}H \cdot K] - P[{\sim}Ba \cdot Ra \mid {\sim}H \cdot K])$

$= P[{\sim}Ba \mid H \cdot K] / (P[{\sim}Ba \mid {\sim}H \cdot K] - (1-p) \cdot P[Ra \mid {\sim}H \cdot K])$

$= (P[{\sim}Ba \mid H \cdot K]/P[Ra \mid {\sim}H \cdot K]) / (q - (1-p))$

$= [q/(q - (1-p))] \cdot P[{\sim}Ba \mid H \cdot K]/P[{\sim}Ba \mid {\sim}H \cdot K]$

> 1 iff $P[{\sim}Ba \mid H \cdot K]/P[{\sim}Ba \mid {\sim}H \cdot K] > [q - (1-p)]/q$.

□

Now we prove Theorem 1. We'll prove it in terms of two distinct names, '*a*' and '*b*', where '*a*' is taken to be an instance of a black raven and '*b*' is taken to be an instance of a non-black non-raven. We do it this way to assure the careful reader that no funny-business is going on when, in the main text, we treat only a single instance '*a*' to see how its turning out to be a black raven compares with its turning out to be a non-black non-raven. To proceed with the treatment in terms of two possibly distinct instances we'll just need to suppose the following: $P[Bb \mid H{\cdot}K] = P[Ba \mid H{\cdot}K]$, $P[Rb \mid H{\cdot}K] = P[Ra \mid H{\cdot}K]$, $P[Bb \mid {\sim}H{\cdot}K] = P[Ba \mid {\sim}H{\cdot}K]$, $P[Rb \mid {\sim}H{\cdot}K] = P[Ra \mid {\sim}H{\cdot}K]$, and $P[Bb \mid Rb{\cdot}{\sim}H{\cdot}K] = P[Ba \mid Ra{\cdot}{\sim}H{\cdot}K]$. The idea is that we have no special knowledge about *b* that permits us to treat it probabilistically any differently than *a* (prior to actually observing it). When *a* and *b* are the same instance, as in the main text, these equalities are tautological.

Theorem 1: Given *Non-triviality*, $q > (1{-}p) > 0$ and

$$\frac{P[Ba{\cdot}Ra \mid H{\cdot}K] \,/\, P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]}{P[{\sim}Bb{\cdot}{\sim}Rb \mid H{\cdot}K] \,/\, P[{\sim}Bb{\cdot}{\sim}Rb \mid {\sim}H{\cdot}K]} = [q-(1-p)]/(p{\cdot}r) > 0.$$

Proof:

To see that $q > (1-p)$, just observe that $q = P[{\sim}Ba \mid {\sim}H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K] =$

$(P[{\sim}Ba{\cdot}Ra \mid {\sim}H{\cdot}K] + P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K]) / P[Ra \mid {\sim}H{\cdot}K] =$

$P[{\sim}Ba \mid Ra{\cdot}{\sim}H{\cdot}K] + (P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K] / P[Ra \mid {\sim}H{\cdot}K]) > (1-p)$, since *Non-triviality* implies $P[{\sim}Ba{\cdot}{\sim}Ra \mid {\sim}H{\cdot}K] > 0$, .

Non-triviality also implies (via Lemma 1) $p = P[Ba \mid Ra{\cdot}{\sim}H{\cdot}K] < 1$; so $0 < 1-p$.

To get the main formula, observe that

$(P[Ba{\cdot}Ra \mid H{\cdot}K] / P[Ba{\cdot}Ra \mid {\sim}H{\cdot}K]) / (P[{\sim}Bb{\cdot}{\sim}Rb \mid H{\cdot}K] / P[{\sim}Bb{\cdot}{\sim}Rb \mid {\sim}H{\cdot}K]) =$

$$(P[Ra \mid H\cdot K] / \{P[Ra \mid {\sim}H\cdot K] \cdot p\}) / (P[{\sim}Bb \mid H\cdot K] / \{P[{\sim}Bb \mid {\sim}H\cdot K] - P[{\sim}Bb\cdot Rb \mid {\sim}H\cdot K]\}) =$$

$$(1/p)\cdot(P[Ra \mid H\cdot K] / P[Ra \mid {\sim}H\cdot K]) \cdot (P[{\sim}Bb \mid {\sim}H\cdot K] - P[Rb \mid {\sim}H\cdot K] \cdot (1-p)) / P[{\sim}Bb \mid H\cdot K] =$$

$$(1/p) \cdot (P[Ra \mid H\cdot K] / P[Ra \mid {\sim}H\cdot K]) \cdot (q - (1-p)) \cdot (P[Rb \mid {\sim}H\cdot K] / P[{\sim}Bb \mid H\cdot K]) =$$

$$(1/p) \cdot (q - (1-p)) \cdot (P[Ra \mid H\cdot K] / P[{\sim}Bb \mid H\cdot K]) \cdot (P[Rb \mid {\sim}H\cdot K] / P[Ra \mid {\sim}H\cdot K]) =$$

$$(1/p) \cdot (q - (1-p))/r.$$

□

Corollary 1: Given *Non-triviality*,

$$\frac{P[Ba\cdot Ra \mid H\cdot K] / P[Ba\cdot Ra \mid {\sim}H\cdot K]}{P[{\sim}Bb\cdot{\sim}Rb \mid H\cdot K] / P[{\sim}Bb\cdot{\sim}Rb \mid {\sim}H\cdot K]} > 1 \text{ if and only if } q - (1-p) > p\cdot r.$$

And, more generally, for any real number s,

$$\frac{P[Ba\cdot Ra \mid H\cdot K] / P[Ba\cdot Ra \mid {\sim}H\cdot K]}{P[{\sim}Ba\cdot{\sim}Ra \mid H\cdot K] / P[{\sim}Ba\cdot{\sim}Ra \mid {\sim}H\cdot K]} = s = [q-(1-p)]/(p\cdot r) > 1 \text{ if and only if}$$

$$[q - (1-p)] = s\cdot p\cdot r > p\cdot r.$$

Proof:

The first biconditional follows from Theorem 1 together with the obvious point that

$[q-(1-p)]/(p\cdot r) > 1$ iff $q - (1-p) > p\cdot r$.

To get the second biconditional just observe that (for any real number s),

$s = [q-(1-p)]/(p\cdot r) > 1$ iff $s\cdot p\cdot r = [q-(1-p)] > p\cdot r$. □

Corollary 2: Given *Non-triviality*, for real number s such that

$$\frac{P[Ba\cdot Ra \mid H\cdot K] / P[Ba\cdot Ra \mid {\sim}H\cdot K]}{P[{\sim}Bb\cdot{\sim}Rb \mid H\cdot K] / P[{\sim}Bb\cdot{\sim}Rb \mid {\sim}H\cdot K]} = s = [q-(1-p)]/(p\cdot r), \text{ we have the following:}$$

(1) $s > (1/p) > 1$ iff $q - (1-p) > r$

(2) $s = (1/p) > 1$ iff $q - (1-p) = r$

(3) $(1/p) > s > 1$ iff $r > q - (1-p) > p \cdot r$.

Proof: Follows easily from Theorem 1.

□

Theorem 2: Given *Non-triviality*, both of the following clauses hold:

(2.1) If $P[\sim Ba \mid H \cdot K] > P[Ra \mid H \cdot K]$ (i.e. if $r > 1$) and

$O[H \mid Ra \cdot K] / O[H \mid \sim Ba \cdot K] > (p + (1-p)/r)$, then

$$\frac{P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid \sim H \cdot K]}{P[\sim Ba \cdot \sim Ra \mid H \cdot K] / P[\sim Ba \cdot \sim Ra \mid \sim H \cdot K]} > 1.$$

(2.2) If $P[\sim Ba \mid H \cdot K] \leq P[Ra \mid H \cdot K]$ (i.e. $r \leq 1$), but either $P[\sim Ba \mid K] > P[Ra \mid K]$

or $P[\sim Ba \mid \sim H \cdot K] > P[Ra \mid \sim H \cdot K]$ (i.e. $q > 1$), then

$$\frac{P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid \sim H \cdot K]}{P[\sim Ba \cdot \sim Ra \mid H \cdot K] / P[\sim Ba \cdot \sim Ra \mid \sim H \cdot K]} > 1.$$

Proof: Assume Non-triviality.

Both parts of the theorem draw on the following observation:

Theorem 1 tells us that $q > (1-p) > 0$ and

$$\frac{P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid \sim H \cdot K]}{P[\sim Bb \cdot \sim Rb \mid H \cdot K] / P[\sim Bb \cdot \sim Rb \mid \sim H \cdot K]} = [q-(1-p)]/(p \cdot r).$$

$$\text{So } \frac{P[Ba \cdot Ra \mid H \cdot K] / P[Ba \cdot Ra \mid \sim H \cdot K]}{P[\sim Bb \cdot \sim Rb \mid H \cdot K] / P[\sim Bb \cdot \sim Rb \mid \sim H \cdot K]} > 1 \text{ iff } [q-(1-p)]/(p \cdot r) > 1 \text{ iff}$$

$q > p \cdot r + (1-p)$ iff $q/r > (p + (1-p)/r)$. We will established each of the two parts of Theorem 2 by showing that their antecedents imply $q/r > (p + (1-p)/r)$.

(2.1) Suppose that $r > 1$ and $O[H \mid Ra \cdot K] / O[H \mid \sim Ba \cdot K] > (p + (1-p)/r)$. Then

$q/r = (P[\sim Ba \mid \sim H{\cdot}K]/P[Ra \mid \sim H{\cdot}K]) \;/\; (P[\sim Ba \mid H{\cdot}K]/P[Ra \mid H{\cdot}K]) =$

$(P[Ra \mid H{\cdot}K]/P[Ra \mid \sim H{\cdot}K]) \;/\; (P[\sim Ba \mid H{\cdot}K]/P[\sim Ba \mid \sim H{\cdot}K]) =$

$$\frac{(P[Ra \mid H{\cdot}K]/P[Ra \mid \sim H{\cdot}K]) \cdot (P[H \mid K]/P[\sim H \mid K])}{(P[\sim Ba \mid H{\cdot}K]/P[\sim Ba \mid \sim H{\cdot}K]) \cdot (P[H \mid K]/P[\sim H \mid K])} = \frac{O[H \mid Ra{\cdot}K]}{O[H \mid \sim Ba{\cdot}K]} > (p + (1-p)/r).$$

(2.2) Suppose $P[\sim Ba \mid H{\cdot}K] \le P[Ra \mid H{\cdot}K]$ (i.e. $r \le 1$), but either $P[\sim Ba \mid K] > P[Ra \mid K]$ or $P[\sim Ba \mid \sim H{\cdot}K] > P[Ra \mid \sim H{\cdot}K]$ (i.e. $q > 1$).

First we show that we must have $q > 1$ in any case. This is shown by reductio, as follows:

Suppose $q \le 1$. Then $P[\sim Ba \mid K] > P[Ra \mid K]$.

So we have $P[\sim Ba \mid \sim H{\cdot}K] \le P[Ra \mid \sim H{\cdot}K]$ (i.e. $q \le 1$) and

$P[\sim Ba \mid H{\cdot}K] \le P[Ra \mid H{\cdot}K]$ (i.e. $r \le 1$). Then

$P[\sim Ba \mid K] = P[\sim Ba \mid H{\cdot}K]\, P[H \mid K] + P[\sim Ba \mid \sim H{\cdot}K]\, P[\sim H \mid K] \le$

$P[Ra \mid H{\cdot}K]\, P[H \mid K] + P[Ra \mid \sim H{\cdot}K]\, P[\sim H \mid K] = P[Ra \mid K] < P[\sim Ba \mid K]$

Contradiction !!!

Thus we have $q > 1$ and $r \le 1$; so $1/r \ge 1$. Then $(p + (1-p)/r) \le p/r + (1-p)/r = 1/r < q/r$.

□

Theorem 3: If the degree to which a non-black non-raven incrementally confirms 'All ravens are black', as measured by the likelihood-ratio, is in the interval

$1 < P[\sim Ba{\cdot}\sim Ra \mid H{\cdot}K] \;/\; P[\sim Ba{\cdot}\sim Ra \mid \sim H{\cdot}K] \le 1+\varepsilon$, for very small $\varepsilon > 0$, then $([q-(1-p)]/q) <$

$P[\sim Ba \mid H{\cdot}K] \;/\; P[\sim Ba \mid \sim H{\cdot}K] \le ([q-(1-p)]/q)\cdot(1+\varepsilon)$.

If instead $(1-\varepsilon) < P[\sim Ba{\cdot}\sim Ra \mid H{\cdot}K] \;/\; P[\sim Ba{\cdot}\sim Ra \mid \sim H{\cdot}K] \le 1$, then

$([q-(1-p)]/q)\cdot(1-\varepsilon) < (P[\sim Ba \mid H{\cdot}K] \;/\; P[\sim Ba \mid \sim H{\cdot}K]) \le ([q-(1-p)]/q)$.

In both cases, for large q, $([q-(1-p)]/q) \approx 1$, so $P[\sim Ba \mid H{\cdot}K] \;/\; P[\sim Ba \mid \sim H{\cdot}K] \approx 1$.

Proof: $P[\sim Ba\cdot\sim Ra \mid H\cdot K] / P[\sim Ba\cdot\sim Ra \mid \sim H\cdot K] = P[\sim Ba \mid H\cdot K] / (P[\sim Ba \mid \sim H\cdot K] - P[\sim Ba\cdot Ra \mid \sim H\cdot K]) = P[\sim Ba \mid H\cdot K] / (P[\sim Ba \mid \sim H\cdot K] - (1-p)\cdot P[Ra \mid \sim H\cdot K]) = (P[\sim Ba \mid H\cdot K] / P[Ra \mid \sim H\cdot K]) / (q - (1-p)) = (P[\sim Ba \mid H\cdot K] / P[\sim Ba \mid \sim H\cdot K]) \cdot q/(q - (1-p))$. So, for $1 < P[\sim Ba\cdot\sim Ra \mid H\cdot K] / P[\sim Ba\cdot\sim Ra \mid \sim H\cdot K] \leq (1+\varepsilon)$, $([q-(1-p)]/q) < (P[\sim Ba \mid H\cdot K] / P[\sim Ba \mid \sim H\cdot K]) \leq ([q-(1-p)]/q)\cdot(1+\varepsilon)$. Also, if $(1-\varepsilon) < P[\sim Ba\cdot\sim Ra \mid H\cdot K] / P[\sim Ba\cdot\sim Ra \mid \sim H\cdot K] \leq 1$, then $([q-(1-p)]/q)\cdot(1-\varepsilon) < (P[\sim Ba \mid H\cdot K] / P[\sim Ba \mid \sim H\cdot K]) \leq [q-(1-p)]/q$.

□

Acknowledgements. We would like to thank the following people for useful conversations about the paradox of confirmation: Luc Bovens, Fabrizio Cariani, Kenny Easwaran, Ted Hailperin, Alan Hájek, Stephan Hartmann, Chris Hitchcock, Colin Howson, Franz Huber, Jim Joyce, Patrick Maher, Chad Mohler, Brad Monton, Mike Titelbaum, and Brian Weatherson. Special thanks to Jan Sprenger, whose critique of the implications of an earlier version of the results in sections 8 and 9 spurred us to make significant improvements.

References

Alexander, H. G. (1958), "The Paradoxes of Confirmation", *British Journal for the Philosophy of Science* 9: 227-233.

Armstrong, D. (1978), *A Theory of Universals*, Cambridge: Cambridge University Press.

Carnap, R. (1950), *Logical Foundations of Probability*. Chicago: Chicago University Press. Second edition 1962.

Carnap, R. (1952), *The Continuum of Inductive Methods*. Chicago: University of Chicago Press.

Carnap, R. (1971), “A Basic System of Inductive Logic, Part I”, in Rudolf Carnap and Richard C. Jeffrey (*eds*.), *Studies in Inductive Logic and Probability*, volume 1. Berkeley: University of California Press.

Carnap, R. (1980), “A Basic System of Inductive Logic, Part II”, in Richard C. Jeffrey (*ed*.), *Studies in Inductive Logic and Probability*, volume 2. Berkeley: University of California Press.

Chihara, C. (1981), “Quine and the Confirmational Paradoxes”, in Peter A. French, Theodore E. Uehling Jr., and Howard K. Wettstein (eds.), *Midwest Studies in Philosophy. Vol. 6: The Foundations of Analytic Philosophy*. Minneapolis: University of Minnesota Press, 425-452.

Earman, J. (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.

Eells, E. (1982), *Rational Decision and Causality*. New York: Cambridge University Press.

Eells E. and B. Fitelson (2000), “Measuring Confirmation and Evidence,” *Journal of Philosoph*y XCVII: 663-672.

Eells, E. and B. Fitelson (2002), “Symmetries and Asymmetries in Evidential Support” *Philosophical Studies 107*: 129-142.

Fitelson, B. (1999), “The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity.” *Philosophy of Science 66*, S362–S378.

Fitelson, B. (2001), *Studies in Bayesian Confirmation Theory*, Ph.D. Dissertation, University of Wisconsin.

Fitelson, B. (2004), “Inductive Logic” forthcoming in *The Philosophy of Science: An Encyclopedia*, Routledge Press, *eds*. Pfeifer, J. and S. Sarkar.

Gaifman, Haim (1979), “Subjective Probability, Natural Predicates and Hempel’s Ravens”, *Erkenntnis* 14: 105-147.

Gemes, Ken (1999), “Carnap-Confirmation, Content-Cutting, and Real Confirmation”, Yale: unpublished essay.

Gibson, L. (1969), “On ‘Ravens and Relevance’ and a Likelihood Solution of the Paradox of Confirmation”, *British Journal for the Philosophy of Science* 20: 75-80.

Good, I. J. (1960), “The Paradox of Confirmation”, *British Journal for the Philosophy of Science 11*: 145-149.

Good, I. J. (1961), “The Paradox of Confirmation (II)”, *British Journal for the Philosophy of Science 12*: 63-64.

Good, I. J. (1967), “The White Shoe is a Red Herring", *British Journal for the Philosophy of Science 17*: 322.

Good, I. J. (1968), “The White Shoe *qua* Herring is Pink", *British Journal for the Philosophy of Science 19*: 156–157.

Goodman, N. (1954), *Fact, Fiction, and Forecast*. London: Athlone Press.

Hempel, C.G. (1945), “Studies in the Logic of Confirmation I & II' *Mind*, 54.

Hempel, C.G. (1967), “The White Shoe: No Red Herring”, *British Journal for the Philosophy of Science 18*: 239–240.

Hesse, M. (1974), *The Structure of Scientific Inference*. London: MacMillan.

Hintikka, J. (1969), “Inductive Independence and the Paradoxes of Confirmation”, in Nicholas Rescher (ed.), *Essays in Honor of Carl G. Hempel: A Tribute on the Occasion of His Sixty-Fifth Birthday*. Dordrecht: Reidel, 24-46.

Hooker, C. A. and D. Stove (1968), “Relevance and the Ravens”, *British Journal for the Philosophy of Science* 18: 305-315.

Horwich, P. (1982), *Probability and Evidence*. New York: Cambridge University Press.

Hosiasson-Lindenbaum, Janina (1940), “On Confirmation", *Journal of Symbolic Logic 5*: 133–148.

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

Humburg, J. (1986), “The Solution of Hempel’s Raven Paradox in Rudolf Carnap’s System of Inductive Logic”, *Erkenntnis* 24: 57-72.

Jardine, R. (1965), “The Resolution of the Confirmation Paradox”, *Australasian Journal of Philosophy* 43: 359-368.

Joyce, J., "Bayes's Theorem", *The Stanford Encyclopedia of Philosophy* (Winter 2003 Edition), Edward N. Zalta (*ed*.), URL = http://plato.stanford.edu/archives/win2003/entries/bayes-theorem/.

Kolmogorov, A. (1956), *Foundations of Probability* (second English ed.). Providence, Rhode Island: AMS Chelsea Publishing.

Mackie, J. L. (1963), “The Paradox of Confirmation”, *British Journal for the Philosophy of Science* 13: 265-277.

Maher, P. (1999), “Inductive Logic and the Ravens Paradox”, *Philosophy of Science* 66: 50-70.

Maher, P. (2004), “Probability Captures the Logic of Scientific Confirmation”, in *Contemporary Debates in the Philosophy of Science*, *ed*. Christopher Hitchcock, Oxford: Blackwell, 69-93.

Nerlich, G. (1964), “Mr. Wilson on the Paradox of Confirmation”, *Australasian Journal of Philosophy* 42: 401-405.

Popper, K. (1954), “Degree of Confirmation”, *British Journal for the Philosophy of Science* 5: 143-149.

Quine, W.V.O. (1969), "Natural Kinds", in: *Ontological Relativity and other Essays*, Columbia University Press, 114-138.

Shoemaker, S. (1980), "Causality and Properties", reprinted in *Identity, Cause and Mind*, Cambridge: Cambridge University Press.

Suppes, P. (1966), "A Bayesian Approach to the Paradoxes of Confirmation", in Jaakko Hintikka and Patrick Suppes (eds.), *Aspects of Inductive Logic*. Amsterdam: North-Holland, 198-207.

Swinburne, R. (1971), "The Paradoxes of Confirmation – a Survey", *American Philosophical Quarterly* 8: 318-330.

Sylvan, R. & R. Nola. (1991), "Confirmation Without Paradoxes," in *Advances in Scientific Philosophy: Essays in Honour of Paul Weingartner*, *eds*. G. Schurz and G.J.W. Dorn. Amsterdam: Rodopi: 5-44.

Vranas, P. (2004), "Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution", *British Journal for the Philosophy of Science 55*: 545–560.

Wilson, P. R. (1964), "A New Approach to the Confirmation Paradox", *Australasian Journal of Philosophy* 42: 393-401.

Woodward, J. (1985), "Critical Review: Horwich on the Ravens, Projectability and Induction", *Philosophical Studies* 47: 409-428.