The Paradox of Confirmation Generalized

[Blinded for Refereeing] *

January 9, 2009

1 Introduction

In this paper I demonstrate how Bayesian confirmation theory can contribute to an epistemology of several metaphysical concepts found at the core of the natural sciences. I then use the results to generalize the Paradox of Confirmation in a way that, I contend, undermines the standard Bayesian solution to the Paradox as well as the more recent refinements proposed in (Howson & Urbach, 2006), (Fitelson, 2007), (Fitelson and Hawthorne, forthcoming), and (Vranas, 2004). The formal results presented serve as yet another illustration of the inadequacy of the standard Bayesian solutions to the Paradox of Confirmation, prompting Bayesians of all stripes to reject the standard solution. As an upshot, this paper also demonstrates that it is imprudent to ignore metaphysical phenomena when constructing a theory of confirmation.

2 Confirming the Metaphysical

It is part of a venerable tradition in the philosophy of science that analyses in their nascent stages are foremost informed by the sciences themselves. In confirmation theory, the rough template adduced from scientific practise is also sharpened by examining simple probability models involving urns, cards, dice and coins; each imposing constraints on the confirmation function. Certainly, the clarity and precision of these austere cases have done much to promote understanding. Nevertheless, analyses that content themselves with the examination of famous episodes from the history of science and simplistic probability models unduly restrict the range of evidence available to them. In this section, I examine a structural constraint on the confirmation function that stems from the often overlooked *metaphysics* of science.

In particular, I will examine the structure of a group of related objects at the core of scientific theory and practise: dispositions, natural laws, counterfactuals, and causation. For the remainder of this paper, I will refer to this loose collection of associated objects as '*the objects of natural analysis*' or, more simply, as ONAs.[1]

*Research generously supported by SSHRC in the form of a standard MGS research grant.

[1] It is worth pointing out that the robust ontological status of such properties is the subject of heated debate. One notable philosopher who has developed an antirealist position within this debate is (Van Fraassen, 1980).

One of the reasons for singling out ONAs when discussing theories of confirmation is that, in addition to being directly relevant to scientific practise, prominent analyses of ONAs entail *several* verifiable consequences. This result imposes constraints on confirmation functions that will be examined and exploited to generalize the Paradox of Confirmation in §3. The following argument establishes how the common structure of ONAs renders them confirmation conducive:

Argument 1. The Argument from Entailment

1. Deterministic ONA statements generally entail their corresponding counterfactual generalizations.

2. Any counterfactual generalization, $\forall x(Ax \,\Box\!\!\rightarrow\, Bx)$, entails the corresponding material generalization $\forall x(Ax \supset Bx)$, (by §5, Theorem 3).

3. Any material generalization, $\forall x(Ax \supset Bx)$, entails its corresponding instances, $\neg A\alpha \vee B\alpha$ (for any given α).

4. If (1), (2) and (3), then (5), (by the transitivity of entailment).

5. ONAs generally entail their corresponding instances.

6. If h entails e, (for non-trivial h, e), then $\mathbb{P}\left[h\middle|e\right] > \mathbb{P}\left[h\right]$, (by §5, Theorem 1).

7. If (5) and (6), then (8).[2]

8. ONAs gain a confirmation boost when their corresponding instances have been *directly* established, (from *modus ponens* on (7) and the conjunction of (5) and (6)).[3] □

From a Bayesian point of view, premise (1) of Argument 1 is perhaps the least trivial, hence it requires the most defence. That is not to say that premise (1) is not without substantive intuitive appeal. If α causes β, or if δ is disposed to Ψ when Φed, then it seems like β would be the case if α were, or if $\Phi\delta$ were the case then $\Psi\delta$ would be the case. The link between ONAs and counterfactuals are so tight that they have spawned various attempts at reductive analyses of the latter in terms of the former. However, even if the connection between the two is not a strict reduction of one in terms of the other, the above argument retains its potency. As long as deterministic ONA statements

---

However, even those who deny the existence of such objects should admit that they do appear in scientific discourse and that they do so in a way that constrains the way in which theorizing is done. For my current purposes it will be sufficient that research programmes be constrained in this way.

[2]This premise is in some need of qualification, for there are certainly some measures of confirmation on which (5) and (6) do not entail (8). Nevertheless, all confirmation functions that have been proposed in the literature (that I am aware of) employ a confirmation function that validates (7). For the most part, the proofs in this paper content themselves with demonstrations involving the difference measure of confirmation $c\left[h, e\right] = \mathbb{P}\left[h\middle|e \cdot k\right] - \mathbb{P}\left[h\middle|k\right]$ that is most frequently assumed in the literature. See (Fitelson, 1999) though for a more detailed discussion of confirmation function choice.

[3]Note that the ‘directly’ part is important here. Learning something that entails an instance of a material generalization is not guaranteed to confirm that generalization without further background constraints. See (Good, 1967) for a compelling Bayesian counterexample to embracing the so-called “special consequence condition” as it applies in this context.

entail certain corresponding counterfactual generalizations *most of the time* it will still be the case that ONAs are often confirmation conducive. In other words, the conclusion that ONAs are confirmable is robust. A lone counterexample to counterfactual theories of deterministic ONAs does not undermine the argument for every ONA. Every counterfactual supporting ONA is an ONA to which the argument applies. So far then, given that most counterexamples to a strict correspondence between counterfactuals and deterministic ONAs came at great labours, are often quite contrived and are generally "special cases", these loose considerations do a great deal to bolster arguments in the spirit of 1 above. Moreover, a close examination of mainstream analyses of ONAs reveals that they all entail counterfactuals for which the argument can be run, providing substantial additional support for premise (1) of Argument 1 [4]

Summing up, the use of ONAs shouldn't be ruled out of "serious" science on the grounds that they cannot be confirmed. ONA statements entail their confirmation conducive material cousins. Hence, by Theorem 1 of §5, the Appendix, ONA statements are also confirmation conducive. The metaphysics of the objects natural analysis yield important information on how the structure of our subjective credence functions should be constrained. Hence, from a Bayesian point of view, the metaphysics of science in part determines how our choice of confirmation function behaves. Let us now turn our attention to how these observations relate to the Paradox of Confirmation.

3 A Generalization of Hempel's Paradox

The Paradox of Confirmation, as it is usually formalized, consists of two seemingly innocuous premises:

(NC) Nicod's Condition, which states that universal generalizations, $\forall x(Ax \supset Bx)$, are confirmed by their positive instances, $(A\alpha \wedge B\alpha)$, for any object α; and

(EQUV) The Equivalence Condition, which states that a piece of evidence e confirms (disconfirms) a hypothesis h iff e confirms (disconfirms) all hypotheses that are logically equivalent to h.

From these two plausible premises follows the paradoxical conclusion that for any predicates A, B and any object α, $(\overline{A\alpha}\,\overline{B\alpha})$ confirms the hypothesis $\forall x(Ax \supset Bx)$.[5]

[4] In (Author, 2009) I argue that there are good reasons for thinking that any analysis of deterministic causation, even ones that do not attempt to reduce causal claims to counterfactual claims, each entail a counterfactual generalization for which my argument holds. However, rehearsal of the analysis given would take us too far afield here. I invite the interested reader to examine the archetypical accounts of causation given in (Lewis, 1986a, 1986b; Hitchcock, 2001; Salmon, 1984; Van Fraassen, 1980) that broadly represent the different analyses of causation given in the literature; in particular, the first two being counterfactual approaches to causation, the third being a representative of manipulability approaches to causation, Salmon's work illustrates the machinery employed by process analyses of causation, finally Van Fraassen's work demonstrates how objective conditional probabilities (and hence probabilistic analyses of causation) invoke the counterfactuals required by my argument.

[5] The symbolic notation in this and the remaining sections is that common to treatments of probability and lends itself well to readability in proof. The symbolization is as follows: sentences of the form $\ulcorner\Phi$ and $\Psi\urcorner$ will be symbolized as $\ulcorner\Phi\,\Psi\urcorner$; $\ulcorner\Phi$ or $\Psi\urcorner$ as $\ulcorner\Phi \vee \Psi\urcorner$; $\ulcorner$If Φ, then $\Psi\urcorner$ as $\ulcorner\Phi \supset \Psi\urcorner$; and $\ulcorner$ not-$\Phi\urcorner$ as $\ulcorner\overline{\Phi}\urcorner$.

The suggested solutions to this problem are multifarious both in scope and in kind. For the moment my focus will only be on Bayesian solutions to the paradox that are of the "canonical" type.[6] These solutions typically accept the argument for the paradoxical conclusion as sound but qualify it in a way that eliminates its bite.[7] In particular, Bayesians usually argue that negative instances ($\overline{A\alpha}\,\overline{B\alpha}$) confirm the usual hypotheses $\forall x(Ax \supset Bx)$ only to a negligeable degree, thus explaining away the intuition that negative instances of a hypothesis do not confirm. To spell out the details using a concrete example let the predicate 'A' be the '...is a raven' predicate, 'B' be the predicate '...is black', and consider the customary hypothesis $h = \forall x(Ax \supset Bx)$, that all ravens are black. If we let $e = (\overline{A\alpha}\,\overline{B\alpha})$, what the Bayesian needs to show is that $c[h, e] = \varepsilon$, for acceptable confirmation function(s), $c[\,,\,]$, and some negligibly small ε. In order to derive the desired result Bayesians usually assume that the number of non-black objects dwarfs the number of ravens and that $\mathbb{P}\left[A\alpha \middle| \overline{B\alpha}\right] = \mathbb{P}[A\alpha]$. The desired result can easily be shown to follow for standard confirmation functions $c[\,,\,]$.[8]

We are now in a position to show that Bayesians who accept the standard solution to the paradox of confirmation are susceptible to a generalized version of the paradox. By the standard Bayesian solution to the paradox of confirmation, many negative instances of any material conditional confirm that conditional. Now let $\mathbb{P}_{\bar{A}\bar{B}}[\,\cdot\,]$ be the probability distribution arrived at by conditionalizing on such a negative instance of some particular universal generalization, (UG) $= \forall x(Ax \supset Bx)$, ε be the amount of additional probability ($\overline{A\alpha}\,\overline{B\alpha}$) lends (UG), and consider the amount of confirmation ($\overline{A\alpha}\,\overline{B\alpha}$) bestows on a corresponding ONA. For concreteness, let the ONA and negative instance in question be such that the standard solution's assumptions apply, i.e. that our credence distributions for the conditional entailed are such that $\mathbb{P}[A\alpha] \ll \mathbb{P}\left[\overline{B\alpha}\right]$ and $\mathbb{P}\left[\overline{B\alpha} \middle| h\right] = \mathbb{P}\left[\overline{B\alpha}\right]$. For example, we could let the ONA be the causal relation that submitting carbon materials to 60 kilobars of pressure at a temperature of 1000°C will cause that material to become a diamond, and let the negative instance be the observation of a brown bookcase.

$$c\left[ONA, \overline{A\alpha}\,\overline{B\alpha}\right] = \mathbb{P}_{\bar{A}\bar{B}}[ONA] - \mathbb{P}[ONA]$$

Applying the law of total probability on the partition created by UG, $\overline{UG}$ to $\mathbb{P}_{\bar{A}\bar{B}}[ONA]$ we obtain

$$\left(\mathbb{P}_{\bar{A}\bar{B}}\left[ONA \middle| UG\right]\mathbb{P}_{\bar{A}\bar{B}}[UG] + \mathbb{P}_{\bar{A}\bar{B}}\left[ONA \middle| \overline{UG}\right]\mathbb{P}_{\bar{A}\bar{B}}\left[\overline{UG}\right]\right) - \mathbb{P}[ONA]\,.$$

[6]This type of solution seems to be the current industry standard. It is discussed in (Howson & Urbach, 2006; Fitelson, 2006; Fitelson & Hawthorne, Forth.; Vranas, 2004).

[7]More precisely, Bayesians accept that the premises hold, *not* universally, but merely in a wide variety of circumstances. They are usually assumed to hold for the "standard" predicates discussed in this paper. Analogously, they appear to hold for all hypotheses examined in the remainder of this paper and it will be assumed that they do in fact hold unless otherwise stated.

[8]See (Vranas, 2004) and (Fitelson, 1999) for discussion, and §5, Proposition 2, of this paper for a demonstration of the desired result for the confirmation function $c[h, e] = \mathbb{P}\left[h \middle| e\right] - \mathbb{P}[h]$. My proof draws on that of (Vranas, 2004) who proves a slightly more general result in a slightly different way. It should be noted that in the same paper Vranas argues convincingly that the independence assumption, $\mathbb{P}\left[\overline{B\alpha} \middle| h\right] = \mathbb{P}\left[\overline{B\alpha}\right]$, used in the standard solution is unfounded. For a reply see (Howson & Urbach, 2006) and for some plausible weakenings of the disputed assumption see (Fitelson & Hawthorne, Forth.).

But since $(ONA \Rightarrow UG)$[9], it follows that $\mathbb{P}_{\bar{A}\bar{B}}\left[ONA \middle| \overline{UG}\right]$ is zero and we have

$$\left(\mathbb{P}_{\bar{A}\bar{B}}\left[ONA \middle| UG\right]\mathbb{P}_{\bar{A}\bar{B}}\left[UG\right] + 0\right) - \mathbb{P}\left[ONA\right].$$

Expanding $\mathbb{P}[ONA]$ using the law of total probability on the same partition and noting again that $\mathbb{P}\left[ONA \middle| \overline{UG}\right] = 0$ our equation can be written as

$$\mathbb{P}_{\bar{A}\bar{B}}\left[ONA \middle| UG\right]\mathbb{P}_{\bar{A}\bar{B}}\left[UG\right] - \mathbb{P}\left[ONA \middle| UG\right]\mathbb{P}\left[UG\right]. \qquad (*)$$

Now for the most part our credence distributions will be such that if $(\overline{A\alpha}\,\overline{B\alpha})$ is relevant to the ONA at all, it will be relevant through its effects on UG, so that $\mathbb{P}_{\bar{A}\bar{B}}\left[ONA \middle| UG\right] = \mathbb{P}\left[ONA \middle| UG\right]$. But even if this is not the case for our personal probabilities, given that $(\overline{A\alpha}\,\overline{B\alpha})$ serves *at most* to eliminate irrelevant members of the sample space, conditionalizing on it will not decrease $\mathbb{P}\left[ONA \middle| UG\right]$, the probability of our ONA given UG. So (*) is greater than or equal to

$$\mathbb{P}\left[ONA \middle| UG\right]\mathbb{P}_{\bar{A}\bar{B}}\left[UG\right] - \mathbb{P}\left[ONA \middle| UG\right]\mathbb{P}\left[UG\right],$$

which by our definitions is

$$\mathbb{P}\left[ONA \middle| UG\right] \cdot (\mathbb{P}\left[UG\right] + \varepsilon) - \mathbb{P}\left[ONA \middle| UG\right]\mathbb{P}\left[UG\right],$$

for $\varepsilon > 0$. By simple algebra, the $\mathbb{P}\left[ONA \middle| UG\right]\mathbb{P}\left[UG\right]$ terms cancel each other out, thus completing the demonstration that $(\overline{A\alpha}\,\overline{B\alpha})$ confirms its related ONA by a degree of at least (the *positive* number)

$$\mathbb{P}\left[ONA \middle| UG\right] \cdot \varepsilon \qquad \text{(C-ONA)}$$

Hence, the standard assumptions that Bayesians make in their solution to the paradox of confirmation entail that the seemingly irrelevant proposition expressed by $(\overline{A\alpha}\,\overline{B\alpha})$ positively confirms any of its corresponding ONAs. According to their solution then, propositions arrived at by the observation of brown bookcases confirm the propositions expressed by statements like ‘If I were you, I would be out dancing’, ‘carbon materials submitted to 60 kilobars of pressure at a temperature of 1000°C have a disposition to become diamonds’, or ‘travelling at speeds in excess of the speed of sound causes a sonic boom’! But the $(\overline{A\alpha}\,\overline{B\alpha})$ propositions arrived at by observing brown bookcases *do not* confirm such statements. While accepting the paradoxical conclusion of Hempel’s argument is mildly disquieting in the case of material universal generalizations, it is far more unsettling, if not outright disturbing, in the case of ONA statements.

3.1 A Counterargument to the Generalized Paradox

In order to evade the generalized paradox above, Bayesians that adopt the standard solution to the paradox might try to generalize their solution along familiar lines. They might attempt to explain away the above results as follows:

[9]See the defence of premise (1), Argument 1 for details.

Argument 2. A Generalization of the Standard Solution.

1. (NC), (EQUIV) and the Bayesian framework as it applies to confirmation are all sound assumptions, (by their strong intuitive appeal).
2. If (1), then (3), (by §5, Proposition 2).
3. The degree of paradoxical confirmation in the canonical case is in fact minute.
4. But if (1) and (3), then (5).
5. The standard solution to the paradox of confirmation has been vindicated.
6. The degree of paradoxical confirmation entailed by (1) in the generalized paradox is also minute, (by Equation C-ONA above, the fact that ε is minute and the fact that $\mathbb{P}\left[ONA \middle| UG\right] < 1$).
7. If (6), then the confirmation bestowed upon ONAs is acceptable, (by the reasoning in premises (1)-(5)).

Unfortunately, none of the premises of this argument hold. Some fail for familiar reasons. For instance, many have denied one or more of the assumptions in (1). It is now commonplace to deny that (NC), the condition that universal statements are confirmed by their instances, holds in full generality. For instance, (Howson & Urbach, 2006, p. 102) paraphrases a convincing counterexample due to (Rosenkrantz, 1977, p. 35) as follows:

> Three people leave a party, each with a hat. The hypothesis that none of the three has his own hat is confirmed, according to Nicod [NC], by the observation that person 1 has person 2's hat and by the observation that person 2 has person 1's hat. But since the hypothesis concerns only three, particular people, the second observation must *refute* the hypothesis, not confirm it.

Objecting to (NC) along different lines (Quine, 1969) argues that evidence phrased in terms of non-natural kinds, like non-blackness or non-ravenity, is not confirmation conducive; thereby blocking the inference to the paradoxical conclusion. The remaining premises are also tenacious. (Vranas, 2004) has argued convincingly that there is a lacuna in the justification for the conditional premise (2) and the "correct" Bayesian framework for confirmation is also a matter of live debate (Fitelson, 1999).[10]

Though I have not seen any denial of premise (4) in the literature, counterexamples to it too may be generated on the grounds that its antecedent lacks the requisite breadth to establish its consequent. The argument in which premise (4) is embedded provides a resolution to the paradox phrased in terms of the canonical predicates '. . . is a raven' and '. . . is black'. This would be unproblematic if the predicates were mere place holders in the argument; however, they are not. Semantic facts about the denotation of the predicates are employed to establish the antecedent of (4) in a way that is not guaranteed to generalize and, in fact, does not generalize. To see the force of

[10] In particular, Vranas argues convincingly that the independence assumption, $\mathbb{P}\left[A\alpha \middle| \overline{B\alpha}\right] = \mathbb{P}\left[A\alpha\right]$, used in the standard Bayesian solution to the Paradox of Confirmation is unfounded.

this worry let us step away from the paradox as framed by ornithology and consider it from the point of view of an entomologist. Let 'A' be the predicate '... is a metapleural gland holder'[11], 'B' be the predicate '... is an ant' and let our domain of interest be the class of insects $\mathcal{I}$. Now assume $\forall x(Ax \supset Bx)$ is the hypothesis h we wish to confirm and let us compare the degree to which an instance ($A\alpha B\alpha$) of h bestows confirmation on h with the degree to which a negative instance ($\overline{A\alpha}\,\overline{B\alpha}$) confirms h. It is an empirical fact that just over half of the insects on the planet observed so far have been metapleural-gland-holding ants, hence $\mathbb{P}[A\alpha B\alpha] > 0.5$ and $\mathbb{P}\left[\overline{A\alpha}\,\overline{B\alpha}\right]$ is small. h is well confirmed, all metapleural-gland-holders observed so far have been ants and all but a few species of ants are metapleural-havers. But now we can define a large class of fairly realistic constraints on our probability measure given the facts about the case at hand for which $\mathbb{P}\left[h \middle| A\alpha B\alpha\right] < \mathbb{P}\left[h \middle| \overline{A\alpha}\,\overline{B\alpha}\right]$[12] and so on most Bayesian measures of confirmation $c[h, A\alpha B\alpha] < c\left[h, \overline{A\alpha}\,\overline{B\alpha}\right]$. But this is absurd. Even if we are willing to accept that seemingly irrelevant observations slightly confirm a hypothesis it is another matter entirely to have a view that entails that seemingly irrelevant observations provide *better* confirmation for a hypothesis than its positive instances for a large class of cases. Note also that the problem is general, it is not due to some peculiarity of the chosen universe of discourse. If we were to formulate a hypothesis in terms of some ubiquitous fundamental matter (like "dark energy" or "dark matter") we would obtain similar results without choosing a restricted domain (i.e. $\mathcal{I}$).[13]

But even if we ignore the faults that have been found with the standard solution and grant the censured premises, the solution to the Generalized Paradox of Confirmation as fleshed out in Argument 2 fails. In what follows I demonstrate why solutions to the Ravens Paradox that are structurally similar to the standard one cannot generally be exported to deal with the generalized paradox.

One can bolster the standard solution to the Ravens Paradox by contemplating some (very) rough and ready observations: that probabilities, which are measure functions, are sensitive to "class size" and that our judgements about "class size" are often in-

[11]Metapleurals are glands thought to be unique to ants that produce and secrete an antibiotic agent that prevents bacteria and fungus from developing both on the ants themselves and within their nests (Hölldobler & Wilson, 1990).

[12]See §5, Proposition 4 for the details.

[13]It is worth mentioning that there is at least one well-known but non-standard "Bayesian" solution to the Ravens Paradox that sidesteps this worry. (Earman, 1992, p. 69-73) shows that given the standard assumptions about the distribution of ravens and non-black things in the world, (i) the probability of the ravens hypothesis h, given as evidence that an object α drawn from the population of ravens is found to be black *is larger than* the probability of h given as evidence that another object β drawn from the population of non-black things turns out to be a non-raven; **exactly when** (ii) the probability that α is a non-black thing given that it was drawn from the population of ravens and $\neg h$ holds *is larger than* the probability that β is a raven given that it drawn from the population of non-black things and $\neg h$ holds. Simplifying slightly, what Earman has argued is that sampling from the sub-population that is more likely to produce falsifiers (the ravens in this case) provides more confirmation than does sampling from the sub-population that is less likely to produce falsifying evidence. Now if you add to this conclusion the premise that falsificationism is a compelling methodology, then it turns out that the ants case above is *not* counterintuitive since, in this case, one is more likely to find a falsifier to the metapleural-gland hypothesis among the non-ants. Though this strategy does sidestep the worry posed by the ants hypothesis, its credibility depends on how compelling one finds the falsificationist methodology. I leave it to the reader to decide one way or the other. No matter the side on which one falls, the solution does not bear on what follows and is in any case distinct from the "standard" solution under examination.

formed by relative frequencies. To illustrate how these ideas apply, consider again the universal hypothesis $h = \forall x(Ax \supset Bx)$. One way to verify h is to round up all of the As and see if they are Bs. This is uncontroversial. Another way would be to round up all of the non-B objects and check if any of them is an A — after all, if we have looked at all the non-B objects without coming across a single A, then anything that is an A must be in the pile of Bs. But this is just a limit case of induction. If we look at all of the non-black objects, save one, it seems that we will still be justified in our conclusion, $\forall x(Ax \supset Bx)$, just a little less justified than we were in the limit case. It also seems that we can repeat this step again and again, each time losing a bit of justification. What all of this seems to suggest is that confirmation should be related to the size of the class examined as compared to the total class size, i.e. related to the relative frequencies observed. This rough argument far from shows that confirmation is equal to the relative frequency but it does bolster the case for the proponent of the standard solution to the Paradox who claims that negative instances of a hypothesis are confirmatory.[14, 15]

Unfortunately, the above explanation does not apply as cleanly to the generalized solution to the Paradox of Confirmation. First of all, observing all of an ONA's (actual) negative instances rarely provides it maximal confirmation. We intuitively think that confirming the underlying connection between an ONA's terms is an essential part of ONA confirmation whereas the connection is less tight in the case of a mere universal statement.

The finitistic character of the standard solution to the "Ravens" Paradox also invites a false sense of comfort. Indeed, the narrowness of the standard example is most evident in light of the Generalized Paradox. The range of an ONA statement can be non-denumerable even when the actual world only contains a finite number of its negative instances.[16] In such domains, the simplistic justification for the standard view via a rough form of a backwards induction fails — the justification of a hypothesis provided by every instance but one need not differ from the justification of a hypothesis provided by every instance but two. Indeed, the phrase 'every instance but ... ' begs elucidation in transfinite cases — probabilities in non-countable domains require careful definition of both field and measure.

The flow of explanation in the standard defence threatens to be turned on its head when expanded to solve the Generalized Paradox. The strong intuitions we have in favour of (NC) and (EQUIV) easily outweigh a "minute" degree of unexpected confirmation for certain universal generalizations.[17] However, since the confirmation of

[14]Solutions along these lines are proposed in the literature as early as (Mackie, 1963). Many since have followed suit.

[15]Unfortunately, the explanation still leaves open the exact relation between observed relative frequencies and confirmation. In particular, it does little to assuage the worry for Bayesian confirmation theory raised in connection with the entomology example. It still would seem odd for a Bayesian to count it amongst the virtues of her theory that sometimes observations that are seemingly irrelevant to a hypothesis provide more confirmation to that hypothesis than its positive instances. This latter point has been discussed at length by (Laetz, 2007).

[16]For instance, current theory seems to underwrite the idea that for any observable state there is a non-denumerable set of possible underlying quantum states. By the counterfactual nature of ONA statements this seems to imply that the merely possible states in near enough worlds may also falsify the hypothesis even if none do at the actual world.

[17]Though, as we have observed, the "minuteness" of the conformation that negative instances afford their hypotheses is a function of both the domain on which they operate and the hypothesis in question. That such

ONAs by their negative instances is *prima facie* much stranger than in the case of universal generalizations, that it is shown to follow from (NC) and (EQUIV) may not be enough to assuage the worry. Firm intuitions may transfer support across a conditional with a mildly surprising conclusion but the *modus ponens* inference risks changing into a case of *modus tollens* when the conclusion becomes more radical. Of course, like many arguments from intuition, these results are not decisive, but that does not mean that they are without substantive pull. A convincing solution to the Paradox of Confirmation should not create as adverse a reaction as the Paradox itself.

4 Getting our Ravens in a Row

In sum, I have argued that the objects of natural analysis are confirmation conducive within the Bayesian confirmation framework. I then put the argument to work, effectively generalizing the Paradox of Confirmation and undermining its "canonical" solution. I take it that what we should take away from the failure of the standard solution is that we should not take the metaphysics of science for granted when constructing our confirmation theories. Metaphysical notions, and especially ONAs, are fertile grounds for testing confirmation theories since their domains are wider than those of their more mundane counterparts.

One might try to sidestep worries caused by esoteric entities by pointing to sciences nominalistic preferences. However, until we cleanly underpin scientific discourse (which seems to blatantly call upon ONAs) in a parsimonious way, it is irresponsible to invoke preferences where arguments are called for. Until the triumph of anti-realism, it seems that we should want to provide a notion of confirmation that is robust enough to handle the basic metaphysical notions that are seemingly explicit parts of both scientific theory and practise. A good place to start then would be by not ignoring metaphysics when formalizing our epistemology.

5 Appendix

Theorem 1. If h entails e, $0 < \mathbb{P}[h]$ and $0 < \mathbb{P}[e] < 1$, then $\mathbb{P}\left[h\middle|e\right] > \mathbb{P}[h]$.

Proof. Assume that h entails e, $\mathbb{P}[h] > 0$ and $0 < \mathbb{P}[e] < 1$. Then

$$\begin{aligned}
\mathbb{P}\left[h\middle|e\right] &> \mathbb{P}\left[h\middle|e\right]\mathbb{P}[e] && \text{(by } \mathbb{P}[e] < 1\text{)} \\
&= \left(\frac{\mathbb{P}[h \wedge e]}{\mathbb{P}[e]}\right)\mathbb{P}[e] && (\textit{df} \text{ of } \mathbb{P}\left[h\middle|e\right], \mathbb{P}[e] > 0) \\
&= \mathbb{P}[h \wedge e] && \text{(by algebra)} \\
&= \mathbb{P}\left[e\middle|h\right]\mathbb{P}[h] && (\textit{df} \text{ of } \mathbb{P}\left[e\middle|h\right], \mathbb{P}[h] > 0 \text{ and algebra}) \\
&= 1 \cdot \mathbb{P}[h] && \text{(by } [h \Rightarrow e]\text{)}
\end{aligned}$$

□

confirmation is minute is a peculiarity of the Ravens example that does not easily generalize.

Proposition 2. For non-trivial $h = \forall x(Ax \supset Bx)$, $e = \overline{A\alpha}\,\overline{B\alpha}$, if $\mathbb{P}[A\alpha] \ll \mathbb{P}\left[\overline{B\alpha}\right]$ and $\mathbb{P}\left[\overline{B\alpha}\middle|h\right] = \mathbb{P}\left[\overline{B\alpha}\right]$, then $\mathbb{P}\left[\forall x(Ax \supset Bx)\middle|\overline{A\alpha}\,\overline{B\alpha}\right] - \mathbb{P}[\forall x(Ax \supset Bx)] = \varepsilon$ for some minute ε.

Proof. Assume that $\mathbb{P}[A\alpha] \ll \mathbb{P}\left[\overline{B\alpha}\right]$ and $\mathbb{P}\left[\overline{B\alpha}\middle|h\right] = \mathbb{P}\left[\overline{B\alpha}\right]$, and that h, e are non-trivial.
-Remark: First note that if $\mathbb{P}[A\alpha] \ll \mathbb{P}\left[\overline{B\alpha}\right]$, then $\mathbb{P}[A\alpha]/\mathbb{P}\left[\overline{B\alpha}\right]$ is minute. Moreover, $\delta = \left(\mathbb{P}\left[A\alpha\overline{B\alpha}\right]/\mathbb{P}\left[\overline{B\alpha}\right]\right) = \mathbb{P}\left[A\alpha\middle|\overline{B\alpha}\right]$ will be *even smaller* than that minute quantity. Then, given our assumptions, it follows that:

$$
\begin{aligned}
\mathbb{P}\left[h\middle|e\right] - \mathbb{P}[e] &= \mathbb{P}[h]\left(\frac{\mathbb{P}\left[\overline{B\alpha}\middle|h\right] \div \mathbb{P}\left[\overline{B\alpha}\right]}{\mathbb{P}\left[\overline{A\alpha}\middle|\overline{B\alpha}\right]} - 1\right) && \text{(by Bayes' Rule and algebra)}\\
&= \mathbb{P}[h]\left(\frac{1}{\mathbb{P}\left[\overline{A\alpha}\middle|\overline{B\alpha}\right]} - 1\right) && \text{(by } \mathbb{P}\left[\overline{B\alpha}\middle|h\right] = \mathbb{P}\left[\overline{B\alpha}\right]\text{)}\\
&= \mathbb{P}[h]\left(\frac{1}{(1-\delta)} - 1\right)^{\dagger} && \text{(by `Remark')}\\
&= \varepsilon && (\text{since } \dagger \approx [\mathbb{P}[h]\cdot(1-1)] = 0)\,. \;\square
\end{aligned}
$$

Theorem 3. $\forall x(Ax \;\square\!\!\rightarrow Bx)$ entails $\forall x(Ax \supset Bx)$

DISCLAIMER In order to sidestep concerns that arise in the metaphysics of modality when considering issues of quantification, the following proof will not be carried out within any particular semantic or syntactic system. Fortunately, accepting the proof will only commit us to the mildest modal constraints. All that we will assume is that (1) when evaluated at a world w, the universal quantifier '$\forall$' ranges over at least the objects of w (whatever 'the objects of w' turns out to mean) and that (2) for any conditional $\ulcorner\phi \;\square\!\!\rightarrow \psi\urcorner$ true at a world w, all worlds which are *ceteris paribus* the same as w and at which $\ulcorner\phi\urcorner$ holds, $\ulcorner\psi\urcorner$ is true there also. Note that this last condition holds even in the weakest standard system of conditional logic C. [18]

proof sketch. For the loose semantic argument, assume that $\forall x(Ax \;\square\!\!\rightarrow Bx)$ holds at a world w. Then by (1) it follows that $(A\alpha \;\square\!\!\rightarrow B\alpha)$ is true there for all α of w. By (2), all worlds which are *ceteris paribus* the same as w and at which $A\alpha$ holds, $B\alpha$ is true there also. Hence, assuming w is *ceteris paribus* the same as w, it follows that if $A\alpha$ holds at w then $B\alpha$ is true at w, so $(A\alpha \supset B\alpha)$ holds at w for all α of w. Thus, $\forall x(Ax \supset Bx)$ holds at w as we wanted to show.

Correspondingly, a syntactic argument would run roughly as follows:

$$
\begin{aligned}
\forall x(Ax \;\square\!\!\rightarrow Bx) &\Rightarrow (A\phi \;\square\!\!\rightarrow B\phi) && \text{(universal instantiation).}\\
&\Rightarrow (A\phi \supset B\phi) && \text{(conditional elimination).}\\
&\Rightarrow \forall x(Ax \supset Bx) && \text{(universal generalization).}
\end{aligned}
$$

Either way, the inference should be uncontroversial. □

[18] See (Nute, 1984; Priest, 2001) for further discussion of standard conditional logics.

Proposition 4. Let h be '$\forall x(Ax \supset Bx)$' and $\mathbb{P}[h] > 0$.
Then if (i) $\mathbb{P}[A\alpha B\alpha] > 0.5$, (ii) $\mathbb{P}\left[\overline{A\alpha}\,\overline{B\alpha}\right] < 0.5$ and (iii) $\mathbb{P}\left[A\alpha \middle| h\right] < (5/3)\mathbb{P}\left[\overline{B\alpha} \middle| h\right]$, in accordance with the facts about metapleural-gland-holders and ants in our example, it follows that $\mathbb{P}\left[h \middle| A\alpha B\alpha\right] < \mathbb{P}\left[h \middle| \overline{A\alpha}\,\overline{B\alpha}\right]$.

Proof. Let (i), (ii) and (iii) hold.
Then $\mathbb{P}\left[A\alpha \middle| h\right] < \mathbb{P}\left[\overline{B\alpha} \middle| h\right] < \left(\mathbb{P}[A\alpha B\alpha] \div \mathbb{P}\left[\overline{A\alpha}\,\overline{B\alpha}\right]\right)\mathbb{P}\left[\overline{B\alpha} \middle| h\right]$, which implies

$$\Rightarrow \frac{\mathbb{P}\left[A\alpha \middle| h\right]\mathbb{P}[h]}{\mathbb{P}[A\alpha B\alpha]} < \frac{\mathbb{P}\left[\overline{B}\alpha \middle| h\right]\mathbb{P}[h]}{\mathbb{P}\left[\overline{A\alpha}\,\overline{B\alpha}\right]}$$

$$\Rightarrow \frac{\mathbb{P}\left[A\alpha B\alpha \middle| h\right]\mathbb{P}[h]}{\mathbb{P}[A\alpha B\alpha]} < \frac{\mathbb{P}\left[\overline{A\alpha}\,\overline{B\alpha} \middle| h\right]\mathbb{P}[h]}{\mathbb{P}\left[\overline{A\alpha}\,\overline{B\alpha}\right]} \qquad \text{(by the content of } h\text{)}$$

$$\Rightarrow \mathbb{P}\left[h \middle| A\alpha B\alpha\right] < \mathbb{P}\left[h \middle| \overline{A\alpha}\,\overline{B\alpha}\right] \qquad \text{(by Bayes Rule).} \qquad \square$$

References

Earman, J. (1992). *Bayes or bust: A critical examination of bayesain confirmation theory*. Cambridge, Massachusets: The MIT Press.

Fitelson, B. (1999, September). The plurality of bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science*, *66*(Proceedings Supplement), S362-S378.

Fitelson, B. (2006, February). The paradox of confirmation. *Philosophy Compass*, *1*. (URL: http://dx.doi.org/10.1111/j.1747-9991.2005.00011.x)

Fitelson, B., & Hawthorne, J. (Forth.). How bayesian confirmation theory handles the paradox of the ravens. In J. Fetzer & E. Eells (Eds.), *The place of probability in science.* Chicago: Open Court.

Good, I. J. (1967). The white shoe is a red herring. *British Journal for the Philosophy of Science*, *17*(4), 322.

Hitchcock, C. (2001, June). The intransitivity of causation revealed in equations and graphs. *The Journal of Philosophy*, *98*(6), 273-299.

Hölldobler, B., & Wilson, E. O. (1990). *The ants*. Cambridge: Belknap Press.

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The bayesian approach*. Peru, Illinois: Open Court.

Laetz, B. (2007, October). *Does the bayesian solution to the paradox of confirmation really support bayesianism?* (Paper presented at the annual Western Canadian Philosophical Association Congress.)

Lewis, D. K. (1986a). Causation. In *Philosophical papers* (Vol. II, p. 159-172). New York: Oxford University Press.

Lewis, D. K. (1986b). Postscripts to 'causation'. In *Philosophical papers* (Vol. II, p. 172-213). New York: Oxford University Press.

Mackie, J. L. (1963). The paradox of confirmation. *British Journal for the Philosophy of Science*, *13*(52), p. 256-277.

Nute, D. (1984). Conditional logic. In D. Gabbay & F. Guenthner (Eds.), *Handbook of philosophical logic: Extensions of classical logic* (Vol. II, chap. 2). Dordrecht: Kluwer.

Priest, G. (2001). *An introduction to non-classical logic*. Cambridge: Cambridge University Press.

Quine, W. V. O. (1969). Natural kinds. In *Ontological relativity and other essays.* New York: Columbia University Press.

Rosenkrantz, R. (1977). *Inference, method, and decision: Towards a bayesian philosophy of science*. Dordrecht: Reidel.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. New Jersey: Princeton University Press.

Van Fraassen, B. C. (1980). *The scientific image*. New York: Clarendon Oxford University Press.

Vranas, P. B. M. (2004, September). Hempel's raven paradox: A lacuna in the standard bayesian solution. *British Journal for the Philosophy of Science*, *55*(3), 545-560.