

## Durham Research Online

---

### Deposited in DRO:

31 March 2020

### Version of attached file:

Published Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Peden, William (2020) 'A conciliatory answer to the paradox of the ravens.', *Journal for general philosophy of science.*, 51 (1). pp. 45-64.

### Further information on publisher's website:

<https://doi.org/10.1007/s10838-019-09485-3>

### Publisher's copyright statement:

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

### Additional information:

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.



# A Conciliatory Answer to the Paradox of the Ravens

William Peden<sup>1,2</sup>

Published online: 27 December 2019  
© The Author(s) 2019

## Abstract

In the Paradox of the Ravens, a set of otherwise intuitive claims about evidence seems to be inconsistent. Most attempts at answering the paradox involve rejecting a member of the set, which seems to require a conflict either with commonsense intuitions or with some of our best confirmation theories. In contrast, I argue that the appearance of an inconsistency is misleading: ‘confirms’ and cognate terms feature a significant ambiguity when applied to universal generalisations. In particular, the claim that some evidence confirms a universal generalisation ordinarily suggests, in part, that the evidence confirms the reliability of predicting that something which satisfies the antecedent will also satisfy the consequent. I distinguish between the familiar relation of confirmation simpliciter and what I shall call ‘predictive confirmation’. I use them to formulate my answer, illustrate it in a very simple probabilistic model, and defend it against objections. I conclude that, once our evidential concepts are sufficiently clarified, there is no sense in which the initial claims are both plausible and inconsistent.

**Keywords** Paradox of the Ravens · Paradox of Confirmation · Inductive logic · Evidence

## 1 Introduction

In the Paradox of the Ravens (PR) a number of plausible claims about confirmation seem to commit us to an excessively broad analysis of evidence, such that discovering non-black non-ravens (or learning sentences reporting them) confirms the hypothesis that ‘All ravens are black’ whenever it also confirms ‘All non-black things are non-ravens’.<sup>1</sup> The challenge is either to reject one of these claims or to learn to live with this paradoxical result. The PR has been a persistent thorn in the side of confirmation theory: if our systems force us to make seemingly bizarre claims about confirming simple universal generalisations, then how can they have authority in more advanced applications? Many answers to the PR

<sup>1</sup> I use ‘single inverted commas’ to refer to sentences or words, but “double inverted commas” to refer to uses of sentences or words.

---

✉ William Peden  
wpeden21@gmail.com

<sup>1</sup> Polytechnic University of the Marche, Ancona, Ancona, Italy

<sup>2</sup> Durham University, Ancona, Durham, UK

either reject the confirmation theories with this result or the intuitions that conflict with it. In contrast, I shall argue that both the apparently paradoxical claims *and* our intuitions are correct, because our intuitions are not about the type of evidential relation that confirmation theorists are explicating. My answer is ‘conciliatory’ in the sense that both those confirmation theorists who accept the ‘paradoxical’ results associated with the PR and those who reject them are both correct, but each group is only correct for one of the two different types of evidential relation. The PR is a misunderstanding caused by an ambiguity in terms like ‘confirms’ and ‘evidence’ when applied to universal generalisations.

Ordinarily, when we discuss positive evidential support for universal generalisations, there is a pragmatic implication that the confirming evidence makes it more reliable to infer that something satisfying the antecedent will also satisfy the consequent. When the evidence does so, it provides what I call ‘predictive confirmation’. This type of evidential support comes apart from confirmation *simpliciter* in the PR and this creates the appearance of paradox.

In Sect. 2, I briefly discuss the PR, its scope, and categorise the literature. In Sect. 3, I define predictive confirmation. I offer my answer to the PR in Sect. 4. I finish by considering some objections in Sect. 5.

## 2 Why Do We Care so Much About Ravens?

### 2.1 The Paradox of the Ravens

Imagine Janina, a logician who is considering the concept of confirmation. Suppose that she has background information B. Let X and Y be some expressions: they could be simple ascriptions of predicates, but they could also be a complex matrix such as a connected series of predicates, an existential statement, a statement using relations like ‘less than’, modal operators, or a combination of these expressions. Consider an X and a Y such that Janina believes the following four claims:

1. ‘All X are Y’ and ‘All  $\neg Y$  are  $\neg X$ ’ are logically equivalent. This claim is sometimes called the “Scientific Laws Condition” (Swinburne 1971, 318).
2. Whatever confirms a hypothesis H, relative to some background information, also confirms any hypothesis that is logically equivalent to H. This is sometimes called the “Equivalence Condition” (Hempel 1945, 12).
3. The hypothesis ‘All  $\neg Y$  are  $\neg X$ ’ is confirmed (relative to background knowledge B) by the information that all of a sample that each satisfy the expression  $\neg Y$  also satisfy the expression  $\neg X$ . This is an instance, for X and Y, of the Nicod Criterion, which states that ‘All  $\Phi$  are  $\Psi$ ’ is always confirmed for *any* expressions  $\Phi$  and  $\Psi$ .<sup>2</sup>
4. The hypothesis ‘All X are Y’ is *not* confirmed, relative to B, by evidence in favour of the hypothesis that all members of a sample that each satisfy the expression  $\neg Y$  also satisfy the expression  $\neg X$ .

<sup>2</sup> Claim (3) can be broadened if we want. For instance, we can extend it to cases when we only know evidence that confirms such a sample report, and thus indirectly confirms ‘All  $\neg Y$  are  $\neg X$ ’. Note that the PR, as I have defined it, does not depend on whether the Nicod Criterion is universally true: the Nicod Criterion might be true for some expressions  $\neg X$  and  $\neg Y$ , but false for some other expressions  $\neg Z$  and  $\neg W$ , then we would still have a problem.

As these four claims are inconsistent, Janina must abandon at least one.<sup>3</sup> For example, if she believes that (a) ‘This non-black thing (such as a white thing) is a non-raven (such as a shoe)’ confirms (b) ‘All non-black things are non-ravens’, but also believes that (a) does not confirm the hypothesis (c) ‘All ravens are black’, then she cannot consistently *also* believe the Scientific Laws Condition and the Equivalence Condition. Summarised in a sentence, my answer to this antimony is that the PR trades on an equivocation, and we should disambiguate ‘confirms’ so that (4) is false on one interpretation and (2) is false on the other interpretation. There is no correct interpretation of (1)–(4) that generates a set of claims that are individually plausible but collectively inconsistent.

## 2.2 Three General Approaches to the Paradox

To provide a detailed map of the huge tangled forest of the PR literature is beyond this article’s scope. However, most answers fall into one of three categories:

- A. The ordinary intuitions are more or less correct: the mistake lies with confirmation theorists who fail to appreciate some crucial condition of evidential support for universal generalisations, like natural kinds or degrees of naturalness. The answers of Quine (1970) and Rinard (2014) are examples.
- B. The ordinary intuitions are incorrect. Typically, they are explained as products of a cognitive illusion, in which people confuse a *very small* degree of confirmation with *no* confirmation. The first attempt was by Hosiasson-Lindenbaum (1940); this approach perhaps reaches its apex with the extremely impressive formal analysis by Fitelson and Hawthorne (2010).
- C. The intuitions are correct, but there is also nothing wrong with the confirmation theories that *seem* to contradict them. According to this answer, there are multiple types of confirmation relations: the intuitions are about a concept of favourable evidence that is distinct from confirmation *simpliciter* that Hempel (and most other confirmation theorists) have attempted to analyse. In the existing versions of this approach to the PR, the additional type of confirmation relation is claimed to be ‘selective’ confirmation, which requires that evidence (1) provides confirmation *simpliciter* for the hypothesis *and* (2) disconfirms a rival, in some suitable sense of ‘rival’. Allegedly, the report ‘This raven is black’ selects in favour of ‘All ravens are black’ against rival hypotheses like ‘All ravens are not black’ (relative to the implicit background knowledge in the PR) whereas ‘This non-black thing is a non-raven’ is consistent with either ‘rival’. This strategy originates with Goodman (1954, 72) while Glymour (1980, 157–160) provides a more sophisticated version.

Each of these approaches has its strengths, but also its challenges. For example, for (A) there is the worry that many confirmation theorists have become quite comfortable with the PR results. Have these confirmation theorists really lost touch with the relevant concepts of naturalness? If this is really such a foundational concept in confirmation theory, it is puzzling that so many philosophers are capable of ignoring its effects in the PR case. Arguably, philosophers’ intuitions are not worth much, but *ceteris paribus* it would

<sup>3</sup> The PR was first discussed in print by Hosiasson-Lindenbaum (1940, 136–141). She attributed the idea to Hempel, who later discussed it in print (Hempel 1945).

be preferable to do justice to them. Furthermore, it is not obvious that the strangeness in the PR is that the evidence fails to give stronger reasons for believing the hypothesis, in some sense of ‘belief’ such as greater numerical credence, acceptability, or some qualitative sense. I agree with (A) that there is something strange about such claims, but it is not clear that it is the same as the strangeness (given our actual background knowledge) of claims like ‘A white shoe evinces the hypothesis that *almost all* ravens are black’ or ‘A red herring evinces that the next raven that I see will be black’, where source of the strangeness is clearly the sense that the evidence fails to give reasons to believe the hypothesis given our background knowledge.

In the case of (B), there is the issue that most people do not have a problem with the claim that non-black non-ravens confirm ‘All non-black things are non-ravens’. Yet any evidence has the same degree of confirmation towards logically equivalent hypotheses. Therefore, if (B) is correct, then people would presumably be just as resistant to the claim that non-black non-ravens confirm ‘All non-black things are non-ravens’ as they are to the claim that non-black non-ravens confirm ‘All ravens are black’. This problem was first identified by Scheffler (1968, 284–285).<sup>4</sup> Similarly, as Fisch (1984, 49) points out, many people think that there is something strange about black ravens confirming ‘All non-black things are non-ravens’. One might respond that people not only tend to ignore small degrees of confirmation, but also tend not to recognise that ‘All ravens are black’ is contrapositable into ‘All non-black things are non-ravens’. However, (1) this is ad hoc, (2) it is implausible that anyone would be ignorant of basic deductive relations *and* have a sense of degrees of confirmation that is analogous to the (many) Bayesian explications of this concept, and (3) knowledge of the contrapositability does not seem, in itself, to remove the strangeness of the PR. (This last point is evinced by philosophers of type (A), because it would be absurd to claim that Quine was ignorant of the contrapositability of universal generalisations!) There is also the concern that, provided that the degree of confirmation for each unobserved non-raven towards ‘All ravens are black’ is a non-zero real number, there will be some quantity of non-black non-ravens (say, the discovery of a trillion stars in a newly detected group of galaxies) which provide *stronger* evidence for ‘All ravens are black’ than discovering a raven, and this seems no less paradoxical than the original PR scenario. Degrees of confirmation are brilliant additions to formal epistemology, but it is unproven that they can explain away the PR.

For (C), the principal problem is identifying a sense of ‘confirmation’ that would be plausible as an interpretation of ‘Non-black non-ravens confirm the hypothesis that all ravens are black’, but also does not simply recreate the paradoxical results. Among the many problems for the selective confirmation approaches to (3), consider the statistical generalisation ‘Just 1% of non-black things are non-ravens’. It seems that, under any plausible analysis of ‘rival hypotheses’, this hypothesis is a rival to ‘All ravens are black’. (For example, the hypotheses cannot both be true given our background information; the statistical generalisation is consistent with what we know; and it also meets Glymour’s other conditions for rivalrous hypotheses.) Yet the statistical generalisation is presumably disconfirmed, relative to the implicit background knowledge in the PR scenario, by evidence

<sup>4</sup> There is a similar problem for those who think that the illusion of paradox is caused by the fact that non-black non-ravens are ‘uninteresting’ confirmations of ‘All ravens are black’, because they mean that instances like  $(Ra \rightarrow Ba)$  are only true because “their antecedents are false” (Schock 1984, 349). Why are non-black non-ravens uninteresting when the universal generalisation is formulated as ‘All ravens are black’, but interesting when formulated as ‘All non-black things are non-ravens’?

such as a sample report that all of a large sample of non-black things are non-ravens. Therefore, discovering such a sample would provide selective confirmation by confirming ‘All ravens are black’ *simpliciter* and disconfirm a rival of this hypothesis, and consequently a report of non-black non-ravens would provide selective confirmation as well as confirmation *simpliciter*.

My criticisms of (A), (B), and (C) sketched above are brief and inconclusive. There are many, many good responses that could be made by supporters of these answers. Indeed, the current debate seems to be at an impasse, with a plethora of objections for each account that might take decades to evaluate and rigorously address. However, they at least provide some challenges that should be met by any new addition to the vast corpus of the PR’s answers.

My answer is a version of (C), but one that is very different from the selective confirmation approach. Akin to them, I shall argue that the key to answering the PR is to distinguish between two different types of confirmation: confirmation *simpliciter* and what I shall call “predictive confirmation”. I shall argue that, when we make this disambiguation, we can come to the conciliatory judgement that the PR is generated by a *misunderstanding* between two fundamentally correct groups rather than a *mistake* by either group. The misunderstanding is caused by the fact that confirmation theorists have focused on confirmation *simpliciter*. (This focus is justifiable, because in the next section it will be apparent that confirmation *simpliciter* has a vastly wider domain than predictive confirmation.) The Equivalence Condition is true for confirmation *simpliciter*, but not predictive confirmation. It is true that discovering non-black non-ravens fails to confirm ‘All ravens are black’ in the predictive sense of ‘confirms’, under the assumptions of the PR, but not the *simpliciter* sense of ‘confirms’. On neither sense of confirmation are (1)–(4) in Sect. 2.1 all intuitively plausible, and therefore there is no real paradox.

### 3 Predictive Confirmation

Before introducing my resolution of the PR, it is necessary to (1) consider some of the pragmatic relations between universal generalisations and predictions, and (2) introduce the concept of predictive confirmation.

#### 3.1 Universal Generalisations and Predictions

If I assert that ‘All ravens are black’, then I seem to be suggesting (in many circumstances) that it is reliable to suppose that, if something was a raven, then it would be black. Yet this hypothesis could be true, given the standard analysis of universal generalisations’ semantics, merely because there are no ravens. In the same way, the hypothesis could be probable, given our evidence, merely because of good evidence that there are no ravens. Even when a universal generalisation is vacuously true but intuitively assertable, such as ‘All ideal gases satisfy the Ideal Gas Law’, we seem to have good reasons to believe that *if* something were an ideal gas (which would require a universe incompatible with our actual physics) then it would satisfy the law.

Pragmatics offers a means of explaining this divergence of assertability and truth/probability: under typical circumstances, the assertion of a universal generalisation suggests that

it is reliable to predict an instance of the consequent given an instance of the antecedent.<sup>5</sup> By analogy, most contemporary logicians agree that ‘P but Q’ has the same semantics as ‘P and Q’ and yet it clearly has different pragmatics, because it typically suggests a contrast between the fact that P and the fact that Q.<sup>6</sup> The exact role that the associated predictions play in reasoning will depend on the particular epistemology; my answer to the PR will be compatible with an extremely wide range of uses of the associated predictions in different theories of reasoning (for example, Bayesian formal epistemologies versus ampliative inference-rule logics like default logics) provided that they have *some* role.

Universal generalisations are not the only way of suggesting such predictions by asserting a general hypothesis. Asserting that ‘Almost all ravens are black’ also suggests, in the absence of defeaters (such as ‘This is an Australian raven and almost all Australian ravens are white’) that it is reliable to predict that some particular raven will be black.<sup>7</sup> While it is easy to give cases where asserting a universal generalisation is not necessary for recommending predictions in this way, it is difficult to think of realistic cases where they are not sufficient. What I say below will not depend on whether they are necessary.

Beyond these observations, there are other pieces of evidence for the notion that universal generalisations have a pragmatic association with predictions of their consequent terms given their antecedent terms. Firstly, contrapositives can have different pragmatics. This pragmatic asymmetry helps resolve some curiosities about universal generalisations. For example, many people have a feeling that ‘All ravens are black’ is ‘about’ ravens, whereas ‘All non-black things are non-ravens’ is ‘about’ non-black things. (Two examples are Wright (1966) and Couvalis (1998, 45). Additionally, Hempel (1945, 17) and Lipton (2007, 79) note that many people seem to have this intuition.) I admit that ‘aboutness’ is anything but precise. Yet this sense could be explained by the idea that the assertion ‘All ravens are black’ typically recommends predicting that something with ravenness will also have blackness, whereas the contrapositive ‘All non-black things are non-ravens’ typically recommends predictions from the *absence* of blackness to the *absence* of ravenhood. In many contexts, the reliability of these predictive policies will differ: if it was true that just 99% of ravens were black, then inferring from something being a raven to its blackness could be a highly reliable policy, yet it would still be possible (though very surprising!) that only a tiny but non-zero percentage of non-black things were non-ravens, and therefore that predicting non-ravenhood from non-blackness would be very unreliable.<sup>8</sup>

The predictions suggested by one formulation of a universal generalisation can also diverge from its contrapositive form when the universal generalisation is probable only because there is a high probability that it is vacuously satisfied. For instance, it is very likely that ‘All planets made of pure platinum are exceptions to the laws of thermodynamics’ is

<sup>5</sup> I am not arguing that these implicatures are part of the semantics of universal generalisations; in fact, as far as possible, I shall remain agnostic about the semantics of these sentences.

<sup>6</sup> There is not always a contrast between the disjuncts, as opposed to some other propositions: Douven (2017, 1543) gives the example of ‘He walks slowly, but he walks’.

<sup>7</sup> The hypothesis that it is reliable to predict instances of Y given instances of X in general does not entail that we should always do so. For instance, we know from the Problem of the Reference Class that the belief of a probabilistic statement about a population might be defeated, in some circumstances, by a belief about a subset of that population. I might believe that it is generally reliable to expect that someone who was awarded a PhD is alive today (due to the explosion of postgraduate studies in the past 70 years) but it would not be rational to have this expectation if I know that the PhD was awarded before 1930.

<sup>8</sup> An inferential policy can be reliable in some contexts and unreliable in others. I shall focus on general reliability in this article, as it is the sense that seems to be particularly important for the PR.

probably true, when this hypothesis is interpreted as a purely extensional hypothesis, but only because pure platinum planets are so improbable. Given our actual background information, we can reliably infer from ‘This is not an exception to the laws of thermodynamics’ to ‘This is not a pure platinum planet’, but we cannot reliably infer from ‘This is a pure platinum planet’ to ‘This is an exception to the laws of thermodynamics’.

Another advantage of postulating a pragmatic connection between universal generalisations and predictions is that it provides a sense in which purely extensional generalisations like (1) ‘All the coins in my pocket are pennies’ and (2) ‘All the coins in my pocket are not pennies’ can be ‘rivals’, even though they are logically consistent according to standard contemporary semantics. Even if we say that they would both be true if my pockets are devoid of coins, we can note that their assertions recommend different predictions: my assertion of (1) would tend to make you expect that, if I reach to my pocket to take out some coins, they will be pennies, whereas my assertion of (2) would tend to make you expect that they will not be pennies. According to my suggested analysis of their pragmatics, hypotheses of the form ‘All X are Y’ and ‘All X are  $\neg$ Y’ are associated with incompatible predictions, even if they are logically compatible.

A third advantage is that, without trying to incorporate modality into the semantics of universal generalisations, we can do justice to this sort of observation: there seems to be something wrong with asserting that ( $\chi$ ) ‘All people who sleep unprotected overnight on the Elephant’s Foot in 2019 go on to live a further 10 years’. (The Elephant’s Foot is an extremely radioactive fused blob of corium that was produced by the Chernobyl disaster in 1986. A few hours of exposure would be swiftly fatal.) The mere fact that it is almost certain that no-one will sleep overnight on the Elephant’s Foot in 2019 seems insufficient for justifiably making such an assertion.<sup>9</sup> In contrast, asserting ( $\eta$ ) ‘All things that do not live a further 10 years after 2019 are not people who sleep overnight on the Elephant’s Foot’ would be rather awkward and not something that we would normally say, but asserting  $\eta$  would lack the strangeness of asserting  $\chi$ . At least part of the contrast might be due to the fact that asserting  $\chi$  suggests some potentially lethal predictions, whereas asserting  $\eta$  would presumably be useless, but would not recommend any unwise predictions.

### 3.2 Confirmation and Universal Generalisations

I shall now argue that the confirmation of universal generalisations is multifaceted: there is both confirmation *simpliciter*<sup>10</sup> and what I shall call *predictive confirmation*. This second form of favourable evidence occurs when the evidence both confirms *simpliciter* a universal generalisation *and* confirms the reliability of making its pragmatically associated predictions. Here is an informal definition of predictive confirmation that is not relativized to any particular confirmation theory:

#### 3.2.1 Predictive Confirmation

E is predictive evidence for a universal generalisation of the form ‘All X are Y’ relative to  $B =_{df}$  (1) E confirms ‘All X are Y’ relative to B and (2) E confirms the prediction that

<sup>9</sup> See Jackson (1987) and Douven (2008) for acute discussions of the assertability of conditionals.

<sup>10</sup> I shall leave this concept unanalysed in this article; a very low-resolution description is that it occurs when the evidence provides a reason to be more confident in the confirmed hypothesis.



Ya relative to  $(B \wedge Xa)$ , where the individual constant  $a$  refers to an otherwise unknown individual,<sup>11</sup> while  $Xa$  and  $Ya$  are the assertions that  $a$  satisfies the expressions  $X$  and  $Y$  respectively.

Thus, if E both confirms ‘All ravens are black’ in the *simpliciter* sense, given our background information, and E confirms the prediction that an unknown individual will be black, given our background information and the postulate that the individual is a raven, then E confirms ‘All ravens are black’ in the predictive sense of confirmation. If necessary,  $a$  could refer to a collection of objects (like a social group or class of chemical elements) rather than a particular individual.

Some clarifications are needed: firstly, I am *not* saying that we should actually infer  $Xa$  without evidence. The exercise of postulating the universal generalisation’s antecedent is imaginative, not inferential: we should ask whether E would support the prediction that  $Ya$  if we knew that  $Xa$ . Secondly, on the nature of B: in simple cases, B is our relevant background information. In cases where  $Xa$  and the relevant background information are inconsistent, B is a set of statements matching our background information except with minimal modifications to achieve consistency with  $Xa$ . This clarification covers both mutual inconsistency and the case where our background information is internally inconsistent.

One might wonder why I include clause (1) in the definition. Carnap (1962, 572–573) defines a similar concept, which he calls “qualified-instance confirmation”, and this concept is similar to predictive confirmation except for clause (1).<sup>12</sup> While predictive confirmation and qualified-instance confirmation are similar, they differ in a way that means that my concept avoids one criticism of Carnap’s concept. As Gower (1997, 221) notes, a hypothesis can have increasing and/or high qualified-instance confirmation even if we accept a counterexample to it. While it is plausible that, if we discover that a large sample of white swans deep in the Amazon rainforest, this new information can confirm that ‘All swans are white’ is a reliable rule-of-thumb, it is not clear that there is a sense of ‘evidence’ in which this information can provide evidence (or ‘confirmation’) for the hypothesis is true. Carnap could answer Gower’s criticism by saying that what he was trying to explicate was exactly this sense of a reliable rule-of-thumb. That response is plausible to me, but it highlights the difference between my explication and Carnap’s: I am trying to explicate cases where people say that evidence does or does not provide evidence *for* a universal generalisation, rather than merely the reliability of the hypothesis as a rule-of-thumb. Nonetheless, I must acknowledge a debt of inspiration to Carnap; predictive confirmation could even be understood as confirmation *simpliciter* plus Carnap’s qualified-instance induction.

I define predictive disconfirmation in an analogous way to predictive confirmation: E is predictive evidence against ‘All X are Y’ relative to B if and only if E disconfirms ‘All X are Y’ relative to B or E disconfirms the prediction that some unknown individual satisfies Y given B and the postulate that it satisfies X. However, I do not yet know of any cases in the philosophy of science where predictive disconfirmation is a useful concept; I define it for the sake of completeness.

<sup>11</sup> In particular,  $a$  must not be mentioned in E or B. Put another way, what we know about  $a$  is only what we know about any arbitrary thing that satisfies X. Whether we imagine that  $a$  is some as-yet-unknown thing or some imagined addition to the universe does not seem to make a significant difference. The former seems more intuitive, whereas the latter is neater in the unusual cases where our background knowledge contains the information that we are already familiar with every instance of X, such as ‘The ravens that I saw yesterday’.

<sup>12</sup> See Zabell (2004, Sect. 6.1) for a contemporary discussion.

For example, assume that  $\forall(x)(X \rightarrow Y)$  is an acceptable formalisation of ‘All X are Y’ and assume the adequacy of the standard Bayesian analysis of confirmation. (On the Bayesian analysis, confirmation is positive probabilistic relevance: E confirms H relative to B if and only if  $P(H | E \wedge B) > P(H | B)$ .) Given those assumptions.

### 3.2.2 Bayesian Predictive Confirmation

E predictively confirms a universal generalisation ‘All X are Y’ relative to background information B =<sub>df</sub> the following are both true:

1.  $P(\forall(x)(X \rightarrow Y) | E \wedge B) > P(\forall(x)(X \rightarrow Y) | B)$ .
2.  $P(Ya | E \wedge B \wedge Xa) > P(Ya | B \wedge Xa)$ .

Informally, if (1) E confirms H given B and (2) E increases the probability of the prediction that some unknown individual satisfies F, given B and the postulate that the individual satisfies G, then E predictively confirms ‘All X are Y’.<sup>13</sup>

With these details filled in, it is possible to give a very simple Bayesian example of where predictive confirmation and confirmation *simpliciter* come apart.<sup>14</sup> Imagine that you are playing game with a friend where you can offer each other bets on the overall distribution of ‘heads’ and ‘tails’ in exactly 10 tosses of a two-sided coin that you both know to be fair. The bets can be offered at any time, though both players must accept them. Suppose that the coin has been tossed 5 times and landed ‘heads’ on each occasion. Knowing this information E provides you some evidence that ‘All 10 coin tosses in the game will land heads’ and thus makes it more rational to accept relatively poor odds that this universal generalisation is true. However, E does not provide predictive confirmation for the universal generalisation. The tosses are independent, and therefore if we suppose that some otherwise unspecified toss *a* is one of the remaining 5 tosses in the game, then the probability that *a* lands heads given E is the same as the prior probability of 0.5. Consequently, on a Bayesian identification of confirmation *simpliciter* with positive probabilistic relevance, the two concepts can come apart.

A fundamental difference between the standard Bayesian definition of confirmation (which I am *not* criticising as such) and predictive confirmation is that the Equivalence Condition holds for the former but not the latter. Consider clause (2) in the Bayesian definition of predictive confirmation. For ‘All F are G’ (using F and G to stand for some particular predicates) this clause requires that E confirms Ga given Fa and B. For ‘All  $\neg$ G are  $\neg$ F’, the clause requires that E confirms  $\neg$ Fa given  $\neg$ Ga and B. Yet there will be many circumstances in which E confirms Ga given Fa and B, but not  $\neg$ Fa given  $\neg$ Ga and B, or vice versa. I shall discuss a simple case in Sect. 4.2.

Predictive confirmation has advantages that are very similar to some of the considerations that I noted towards the end of Sect. 3.1. Firstly, since the same evidence cannot confirm both the prediction that Ya and the prediction that  $\neg$ Ya, relative to the same background knowledge and assumption of Xa, it follows that the same evidence cannot predictively confirm both that ‘All X are Y’ and that ‘All X are  $\neg$ Y’. In *this* sense of ‘evidence’,

<sup>13</sup> Note the similarity here to the Ramsey Test for conditionals (Ramsey 1990). In particular, clause (2) should be understood as similar to the hypothetical steps taken for simple conditionals on Ramsey’s view.

<sup>14</sup> I owe this example to an anonymous referee. See also an example by Dretske (1977).

there cannot be evidence that supports both ‘All phlogiston is radioactive’ and ‘All phlogiston is not radioactive’. Secondly, evidence that no-one will visit the Elephant’s Foot in 2019 confirms ‘All people who sleep unprotected overnight on the Elephant’s Foot in 2019 go on to live a further 10 years’ in the *simpliciter* sense, but not in the predictive sense, because the evidence fails to confirm the prediction that a person who slept unprotected overnight on the Elephant’s Foot in 2019 would go on to live a further 10 years and thus satisfy the second clause requirement for predictive confirmation. Finally, there is a significant sense in which hypotheses like ‘All Higgs bosons are electrically charged’ are genuinely about their antecedents, even if they are logically equivalent to hypotheses with different antecedents. This hypothesis is only predictively confirmed by evidence that favours the prediction that an unknown Higgs boson would be electrically charged, and such predictions are not contrapositable.

Although predictive confirmation is a novel idea, at least as I have defined it, and it does not seem to presuppose any particularly controversial theses in the philosophy of language, there are affinities between my notion and some recent work on conditionals. According to the inferentialist theory of conditionals, an utterance of ‘If  $P$ , then  $Q$ ’ is true if and only if (1)  $P$  is evidentially relevant to  $Q$  given the utterer’s background knowledge<sup>15</sup> and (2)  $P$  is consistent with those background knowledge or else is evidence for  $Q$  in the absence of relevant background beliefs (Krzyżanowska et al. 2013; 2014; Douven 2017; Douven et al. 2018). This idea is not much younger than Western philosophy (something like it was apparently proposed by Chrysippus) but unlike some earlier versions of the same idea, the inferential connection does not have to be deductive. This approach is logically independent of my own: it is a thesis about the semantics of conditionals and inferentialists primarily discuss unquantified conditionals, whereas I am concerned with the *pragmatics* of *universally quantified* conditionals. However, we are motivated by similar problems and both make use of inferential notions in our analyses. Much of the evidence I cite for my hypotheses regarding confirmation could also be cited as evidence for the inferentialist analysis and vice versa.<sup>16</sup>

It is perhaps already clear how predictive confirmation will help with the PR. Before going on to make this point in detail, I shall close this section by emphasising that I think that both predictive confirmation and confirmation *simpliciter* are legitimate senses of the claim that a universal generalisation is confirmed. However, predictive confirmation is apparently the typical sense outside of formal epistemology.

## 4 The Resolution of the Raven Paradox

### 4.1 Predictive Confirmation and the Paradox

Let us now return to the PR. Uncontroversially, evidence of  $\neg Y$ ’s that are  $\neg X$  need not confirm the prediction that  $Y$  is true of an unknown individual  $a$ , relative to some background information and the assumption that  $Xa$ , *even though* (by supposition) the evidence

<sup>15</sup> Unlike many uses in the philosophy of science, here ‘background knowledge’ is being used with success grammar for ‘knowledge’. In particular, such background knowledge is *true*.

<sup>16</sup> The inferentialist analysis also has serious consequences for the scope of *modus ponens* as a valid deductive principle (Krzyżanowska et al. 2014, 779–780). My proposal is independent of the scope of *modus ponens*.

confirms the prediction that  $\neg Xa$ , relative to that background information and the assumption that  $\neg Ya$ . Put simply, the same evidence might support the reliability of predictions from instances  $\neg Y$  to  $\neg X$ , without also supporting the reliability of predictions from instances from  $X$  to  $Y$ . Evidence that ‘This non-raven is non-black’ could predictively confirm ‘All non-black things are non-ravens’ without *also* predictively confirming ‘All ravens are black’. The same is also true for evidence of black ravens, which could predictively confirm ‘All ravens are black’ and yet not predictively confirm ‘All non-black things are non-ravens’. In the predictive sense of confirmation, the puzzling PR scenario does not occur.

The PR is simply due to a misunderstanding: confirmation theorists have (justifiably) focused on confirmation *simpliciter*, but our talk about evidential relations is subtle and complex, and the ordinary way of interpreting assertions of the form ‘E is evidence that all X are Y’ is that they are claims about predictive confirmation. Thus, the claim ‘Discovering the existence of my partner’s pair of white shoes provides me evidence for the hypothesis that all ravens are black’ ordinarily sounds like an assertion that is obviously false, assuming the implicit background information, because (in that context) white shoes provide no support for the prediction that an unknown raven would be black. Contrariwise, the claim can seem unparadoxical if one is sufficiently clear that confirmation *simpliciter* (as analysed by a theory like Hempel’s or standard Bayesianism) is the subject of the assertion: for instance, there are probability distributions in which ‘All ravens are black’ is more probable relative to the total evidence after the discovery of some non-black non-ravens, so that arguably we can be more confident in the hypothesis. Once we have disambiguated terms such as ‘evidence for’ or ‘confirms’, we can see that there is a sense in which the commonsense intuitions truly apply and a sense in which they do not apply. It is the latter sense that confirmation theorists are focusing on, and thus there is no fundamental conflict, except among those who extend either sense to where they do not apply.

It is worth making some clarificatory points about my answer. Firstly, I am not claiming that natural language universal generalisations are consistent with counterexamples, nor am I claiming that they are really statistical generalisations or ambiguous with statistical generalisations. I have argued that an assertion such as ‘All panther mushrooms are poisonous’ and an assertion such as ‘Almost all panther mushrooms are poisonous’ have very similar pragmatic roles, but I am not claiming that their semantics are identical or even similar. Secondly, I am not denying that universal generalisations can be confirmed. To the contrary, my answer to the PR depends not only on the possibility that they can be confirmed in the *simpliciter* sense of standard confirmation theory, but *also* in the predictive sense.

Predictive confirmation can come from observing instances of a universal generalisation, as in the case of observing black ravens, but this is not the only possible source of predictive confirmation. A scientist might be investigating the hypothesis that ‘All chromium vaporizes at approximately 347 kilojoules per mole under standard laboratory conditions’, but she might not be in a position to accept that the subject and predicate terms of the hypothesis have been satisfied given her instrument’s readings. Nonetheless, her evidence might confirm that a sample vaporized under those conditions, and with suitable background information thereby confirm the hypothesis.

Relative to some background information, it *can* be the case that statements of the form  $(\neg Xb \wedge \neg Xb)$  provide evidence for the prediction that  $Ya$ , given the postulate that  $Xa$ , such that it predictively confirms statements of the form ‘All X are Y’.<sup>17</sup> For example, imagine

<sup>17</sup> Additionally, relative to some background knowledge, instances of things that are X and Y can fail to provide predictive evidence for ‘All X are Y’, because they disconfirm the universal generalisation and perhaps also the predictive component. To adapt an example from Franklin (2001, 287) no human is known to

that you encounter an Amazonian tribe whose language is largely unknown to you. They seem to be either describing a white raven or a grey parrot, but the language barrier creates difficulties in interpreting their observation reports. This might provide you with evidence against the prediction that some unknown postulated raven (not the bird they are describing) is black. Suppose that, after clarification, you discover that they are referring to a grey parrot. You have learned that something is a non-black non-raven, and it is possible that discovering non-black non-ravens confirms 'All ravens are black' relative to your background knowledge, as in the standard PR scenario. However, it might also confirm the prediction that some unknown raven is black, because it might have seemed relatively likely that there was a white raven (disconfirming your belief of a 100% frequency of blackness in the set of raven) and this possibility was closed-off by discovering that the bird was a grey parrot. In the predictive sense, as well as the *simpliciter* sense, 'That parrot is grey' has confirmed that 'All ravens are black'. Such examples have become standard in the PR literature, and my analysis of predictive confirmation is consistent with their possibility.

Finally, one auxiliary advantage of predictive confirmation is that it provides a type of evidential support that vindicates the intuition that 'All ravens are black' and 'Just 99% of ravens are black' have similar sets of possible confirming evidence-statements. As I said in Sect. 2.2, statistical generalisations are not contrapositable, which is why the PR does not occur for them. Similarly, the conditional predictions ('Given X, expect Y') suggested by universal generalisations are not contrapositable. For predictive confirmation, there is a sense in which both hypotheses are about ravens, but this is due to the formulation of 'All ravens are black' and the pragmatics of this formulation, rather than the semantics of the hypothesis.

At the heart of my resolution is the fact that, while the Equivalence Condition (condition (2) in Sect. 2.1) is a very plausible criterion for any analysis of confirmation *simpliciter*, it need not be true for every sort of evidential relation. For predictive confirmation, the pragmatics of the confirmed or disconfirmed hypotheses are relevant to their relation towards the evidence, and two statements with the same semantics can differ in their pragmatics. Similarly, there is a sense in which it is strange to say that P confirms 'P but Q' relative to B, when no contrast between P and Q is suggested by either B or  $(P \wedge B)$ , even if P clearly confirms the logically equivalent 'P and Q' relative to B. There is no interpretation of the claims in Sect. 2.1 on which the Equivalence Condition is true and yet it is perplexing that reports of non-black non-ravens would confirm 'All ravens are black.'

## 4.2 A Probabilistic Illustration

My discussion in the preceding section was informal, and some readers might legitimately desire a formal illustration of how predictive confirmation avoids the PR. There are two points I shall make: firstly, that confirmation *simpliciter* and predictive confirmation can come apart; secondly, that predictive confirmation does not satisfy the Equivalence Condition, and therefore it is possible for reports of non-black non-ravens to predictively confirm 'All ravens are black', but not 'All non-black things are non-ravens'. Thus, in the set of

---

Footnote 17 (continued)

be taller than 12 feet tall, but if we discovered the skeletons of a group of prehistoric humans, among whom all the adults were between 11.5 feet tall and 11.9 feet tall, we have apparently confirmed that some human has been over 12 feet tall. This is because the height of our sample of skeletons indicates that there were some unsampled of this group who were over 12 feet tall.

claims I outlined in Sect. 2.1, they are all true for predictive confirmation except the Equivalence Condition; contrariwise, for confirmation *simpliciter*, reports of black ravens really do confirm ‘All ravens are black’ (given the implicit assumptions of the PR, the reports really do increase the probability of the hypothesis) and this only seems strange because we tend to refer to predictive confirmation when making evidential claims about universal generalisations.

I shall consider a very simple example with a very small domain, consisting of two objects *a* and *b*, characterised by two logically independent predicates S and G. My example below can be considered in the abstract, but if you would like to imagine circumstances where we would use such a probability model, imagine that *a* and *b* are two lottery balls that have just been drawn by a machine from a vat behind a screen. Let S and G be the predicates ‘small’ and ‘green’ respectively. Initially, you cannot see either ball, but you will first be shown ball *b*, and then shown ball *a*. You know some facts about the machine, which lead you to believe, in broad terms, that the features of *b* are a very good guide to the features of *a* when *b* is small and green. To a lesser extent, the features of *b* are a good guide when *b* is small and not green. Otherwise, the features of *b* are not helpful. Let B be your relevant background knowledge. To simplify, imagine that P(B)=1. In detail, suppose that your background information results in the following probabilities for the possible circumstances:

- (1)  $P(Sa \wedge Ga \wedge Sb \wedge Gb \wedge B) = \frac{15}{32}$
- (2)  $P(\neg Sa \wedge Ga \wedge Sb \wedge Gb \wedge B) = \frac{1}{32}$
- (3)  $P(Sa \wedge \neg Ga \wedge Sb \wedge Gb \wedge B) = \frac{1}{32}$
- (4)  $P(Sa \wedge Ga \wedge \neg Sb \wedge Gb \wedge B) = \frac{1}{32}$
- (5)  $P(Sa \wedge Ga \wedge Sb \wedge \neg Gb \wedge B) = \frac{1}{32}$
- (6)  $P(\neg Sa \wedge \neg Ga \wedge Sb \wedge Gb \wedge B) = \frac{1}{32}$
- (7)  $P(\neg Sa \wedge Ga \wedge \neg Sb \wedge Gb \wedge B) = \frac{1}{32}$
- (8)  $P(\neg Sa \wedge Ga \wedge Sb \wedge \neg Gb \wedge B) = \frac{1}{32}$
- (9)  $P(Sa \wedge \neg Ga \wedge \neg Sb \wedge Gb \wedge B) = \frac{1}{32}$
- (10)  $P(Sa \wedge \neg Ga \wedge Sb \wedge \neg Gb \wedge B) = \frac{1}{32}$
- (11)  $P(Sa \wedge Ga \wedge \neg Sb \wedge \neg Gb \wedge B) = \frac{1}{32}$
- (12)  $P(Sa \wedge \neg Ga \wedge Sb \wedge \neg Gb \wedge B) = \frac{1}{32}$
- (13)  $P(\neg Sa \wedge Ga \wedge \neg Sb \wedge \neg Gb \wedge B) = \frac{1}{32}$
- (14)  $P(\neg Sa \wedge \neg Ga \wedge Sb \wedge \neg Gb \wedge B) = \frac{1}{32}$
- (15)  $P(\neg Sa \wedge \neg Ga \wedge \neg Sb \wedge Gb \wedge B) = \frac{1}{32}$
- (16)  $P(\neg Sa \wedge \neg Ga \wedge \neg Sb \wedge \neg Gb \wedge B) = \frac{1}{32}$

I shall begin by demonstrating that statement  $(\neg Sb \wedge \neg Gb)$  confirms *simpliciter* ‘All S are G’ in this probability distribution. Let H be ‘All S are G’. The main intuition behind the calculations in this paragraph is that if  $(\neg Sb \wedge \neg Gb \wedge B)$  is true, then there are four equiprobable cases; in three of them, H is true; and this exceeds the probability of H given B alone. Firstly, since P(B)=1, it follows that  $P(H | B) = P(H)$ , and this is the probability that everything is  $\neg S$  or G. That is equal to the sum of the probabilities in (1), (2), (4), (6), (7), (11), (13), (15), and (16), which is  $\frac{23}{32} = 0.71875$ . Secondly, the probability of  $(\neg Sb \wedge \neg Gb \wedge B)$  is the sum of the probabilities in (11), (12), (13), and (16), which is  $\frac{4}{32}$ . Finally,  $(H \wedge \neg Sb \wedge \neg Gb \wedge B)$  is true in the possibilities in (11), (13), and (16), whose probabilities sum to  $\frac{3}{32}$ . The conditional probability of H given  $(\neg Sb \wedge \neg Gb \wedge B)$  is  $P(H | \neg Sb \wedge \neg Gb \wedge B) = \frac{3/32}{4/32} = \frac{3}{4} = 0.75$ . Since this probability is greater than  $P(H | B) = 0.71875$ , it follows that  $(\neg Sb \wedge \neg Gb)$  confirms *simpliciter* H relative to B.

Yet  $(\neg Sb \wedge \neg Gb)$  does not predictively confirm H relative to B. The key feature of the probability distribution behind the calculations in this paragraph is that, given B and the assumption of Sa, the prediction of Ga is initially somewhat more likely than not; however, learning  $(\neg Sb \wedge \neg Gb)$  reduces the possibilities to two equiprobable cases, and Ga is only true in one of these, so that Ga is no longer more likely than not. Firstly, the conditional probability of Ga given  $(B \wedge Sa)$  is the sum of the probabilities in (1), (4), (5), and (11),

which is  $\frac{18}{32}=0.5625$ . Secondly,  $P(\neg Sb \wedge \neg Gb \wedge B \wedge Sa)$  is the sum of the probabilities in (11) and (12), which is  $\frac{2}{32}$ . Finally, the value of  $P(Ga \wedge \neg Sb \wedge \neg Gb \wedge B \wedge Sa)$  is given in (11), which is  $\frac{1}{32}$ . Therefore,  $P(Ga | \neg Sb \wedge \neg Ga \wedge B \wedge Sa) = \frac{P(Ga \wedge \neg Sb \wedge \neg Gb \wedge B \wedge Sa)}{P(\neg Sb \wedge \neg Gb \wedge B \wedge Sa)} = \frac{1/32}{2/32} = \frac{1}{2} = 0.5$ , which is less than  $P(Ga | B \wedge Sa) = 0.5625$ . Far from confirming the prediction in question,  $(\neg Sb \wedge \neg Gb)$  *disconfirms* it.

One might worry that this might be an excessively peculiar probability distribution. In particular, one might wonder if this is a probability distribution in which  $(Sb \wedge Gb)$  does not predictively confirm H, so that it is not a 'normal' inductive probability distribution. One could then worry that, even though what I have said in the previous paragraphs is true, I have not proven that my points could hold when 'All ravens are black' is confirmed by discovering black ravens. This worry is unfounded, because  $(Sb \wedge Gb)$  *does* predictively confirm H in this probability distribution. The basic idea is that  $(Sb \wedge Gb)$  is antecedently expected to be a very good indicator of the features of *a*, and if this indication is correct, then H is true. Firstly,  $P(H \wedge Sb \wedge Gb \wedge B)$  is the sum of the probabilities in (1), (2), and (6), which is  $\frac{17}{32}$ . Secondly,  $P(Sb \wedge Gb \wedge B)$  is the sum of the probabilities in (1), (2), (3), and (6), which is  $\frac{18}{32}$ . Therefore,  $P(H | Sb \wedge Gb \wedge B) = \frac{P(H \wedge Sb \wedge Gb \wedge B)}{P(Sb \wedge Gb \wedge B)} = \frac{17/32}{18/32} = \left(\frac{17}{32}\right) \left(\frac{32}{18}\right) = \frac{544}{576} = 0.9444$  (to 4 decimal places). In this probability distribution, learning  $(Sb \wedge Gb)$  provides much stronger confirmation *simpliciter* for H than learning  $(\neg Sb \wedge \neg Gb)$ . It also provides the predictive component of predictive confirmation. As noted in the previous paragraph,  $P(Ga | B \wedge Sa) = 0.5625$ . The value of  $P(Sb \wedge Gb \wedge B \wedge Sa)$  is the sum of the probabilities in (1) and (3), which is  $\frac{16}{32}$ . Finally,  $P(Ga \wedge Sb \wedge Gb \wedge B \wedge Sa)$  is given in (1), which is  $\frac{15}{32}$ . Therefore,  $P(Ga | Sb \wedge Gb \wedge B \wedge Sa) = \frac{P(Ga \wedge Sb \wedge Gb \wedge B \wedge Sa)}{P(Sb \wedge Gb \wedge B \wedge Sa)} = \frac{15/32}{16/32} = \frac{15}{16} = 0.9375$ , which is greater than  $P(Ga | B \wedge Sa) = 0.5625$ . Thus,  $(Sb \wedge Gb)$  confirms Ga relative to B and the assumption that Sa, and thereby satisfies the predictive component of predictive confirmation as well as the confirmation *simpliciter* component.

To close, I shall use this probability distribution to exemplify one of my key claims: that predictive confirmation does not satisfy the Equivalence Condition. We have already seen that  $(\neg Sb \wedge \neg Gb)$  provides confirmation *simpliciter* for 'All  $\neg G$  are  $\neg S$ ', since it provides confirmation *simpliciter* for the logically equivalent 'All G are S' and Bayesian confirmation satisfies the Equivalence Condition. Now, I need to prove that it also provides the predictive component. Firstly,  $P(\neg Sa | B \wedge \neg Ga)$  is the sum of the probabilities in (6), (14), (15), and (16), which is  $\frac{4}{32} = 0.125$ . Secondly,  $P(\neg Sb \wedge \neg Gb \wedge B \wedge \neg Ga)$  is the sum of the probabilities in (12) and (16), which is  $\frac{2}{32}$ . Finally,  $P(\neg Sa \wedge \neg Sb \wedge \neg Gb \wedge B \wedge \neg Ga)$  is the probability in (16), which is  $\frac{1}{32}$ . Therefore,  $P(\neg Sa | \neg Sb \wedge \neg Gb \wedge B \wedge \neg Ga) = \frac{P(\neg Sa \wedge \neg Sb \wedge \neg Gb \wedge B \wedge \neg Ga)}{P(\neg Sb \wedge \neg Gb \wedge B \wedge \neg Ga)} = \frac{1/32}{2/32} = \frac{1}{2} = 0.5$ , which is greater than  $P(\neg Sa | B \wedge \neg Ga) = 0.125$ . Thus we can see that, in the probability distribution that I have described,  $(\neg Sb \wedge \neg Gb)$  predictively confirms 'All  $\neg G$  are  $\neg S$ ', even though it does not predictively confirm the logically equivalent hypothesis 'All S are G', because they are pragmatically associated with different predictions. Although confirmation *simpliciter* satisfies the Equivalence Condition, this probability distribution illustrates how predictive confirmation does not, and this is the principal formal feature of predictive confirmation that I need for my answer to the PR.

Of course, this distribution concerns a very artificial case, though its simplicity and the choice of probabilities makes easy to see the precise probabilities in question. For a more realistic example, consider the Elephant's Foot hypothesis that I discussed earlier: it would be misleading, in normal circumstances, to say  $(\Phi)$  'The recent discovery of that galaxy is evidence that all people who sleep overnight on the Elephant's Foot in 2019 live a further

10 years'. On the Bayesian version of my answer to the PR, the statement  $\Phi$  is misleading because it suggests to the listener that the astronomical discovery makes the prediction that if someone did sleep overnight on the Elephant's Foot in 2019, they would live a further 10 years, into a more probable prediction. Clearly, this probabilistic relation does not hold for our actual credences. Thus,  $\Phi$  sounds like it is about predictive confirmation, when actually  $\Phi$  is at best only true for confirmation *simpliciter*.

In the ravens case, the same sorts of considerations apply, even though the domain is obviously far larger than in my toy example. Suppose that evidence of non-black non-ravens decreases the probability that objects are ravens and that this decrease overpowers the effect of *increasing* the probability that objects are non-black. (This latter effect could be very small.) It will then confirm 'All ravens are black'. Unlike the case of black ravens, the confirmation comes from providing evidence that ravens are rare, rather than providing evidence that the relative frequency of blackness among ravens is 100%. However, suppose also that it increases the conditional probability, given our background knowledge and the assumption that some object is a raven, that the raven will *not* be black. This is possible, because the increase in probability that objects are non-black will still be present, but the decrease in the probability of ravens will no longer apply. In other words, the evidence has increased the probability that *if* the postulated object was a raven, then it would be non-black. The evidence therefore does not satisfy the predictive component of predictive confirmation, and thus does not predictively confirm 'All ravens are black'.<sup>18</sup>

In either toy examples or more realistic cases, Bayesians can disentangle these two types of evidence and accommodate both the commonsense intuitions in the PR and those philosophers who have been led, by their analyses of confirmation *simpliciter*, to accept what *seem* to be the opposite of the commonsense intuitions.

### 4.3 Comparison with Alternatives

Firstly, unlike approach (A), my answer requires no clash with confirmation theorists like Hempel and most Bayesians. Of course, if the latter group were to insist that confirmation *simpliciter* was the *only* legitimate sense of terms like 'is evidence for', then there would be a clash. However, I know of no reason why confirmation must be a unitary concept in natural language. It would be convenient, but that is no reason to deny predictive confirmation, because natural language is not obliged to be philosophically convenient. Perhaps natural kinds and degrees of naturalness are essential parts of the philosophy of evidence, but if my answer is correct, they are inessential to resolving the PR. Finally, my answer does not require that, in the PR scenario, the evidence fails to probabilify (or otherwise confirm *simpliciter*) 'All ravens are black'.

Unlike approach (B) in Sect. 2.2, my answer does not entail any mistake by those who believe (4) in Sect. 2.1. They are not ignoring very small degrees of confirmation *simpliciter* for 'All ravens are black', while also somehow not ignoring small degrees of confirmation *simpliciter* for the logically equivalent 'All non-black things are non-ravens'. Their intuitions are not about confirmation *simpliciter* at all, except if they extend them beyond predictive confirmation to where they do not belong. Since people's intuitions are fine, when in their proper place, there is nothing to explain away by reference to degrees of

<sup>18</sup> The evidence does not provide predictive disconfirmation, because it does not disconfirm *simpliciter* the hypothesis 'All ravens are black'.



confirmation. My answer is also consistent with the intuitions of those who find it puzzling that a black raven could confirm ‘All non-black things are non-ravens’. As for large numbers of non-black non-ravens, they will provide no confirmation in the predictive sense for ‘All ravens are black’ unless they combine with the relevant background knowledge to confirm the prediction that an otherwise unknown raven would be black.

My answer falls under approach (C) in my classification of answers in Sect. 2.3. We agree that there is a sense of ‘evidence’ where (under typical background assumptions) reports of black ravens provide evidence for ‘All ravens are black’ but non-black non-ravens do not. We also agree that the PR is fundamentally the product of the ambiguity of the notion of ‘evidence’, and that the paradox is dissolved once we clarify the different senses of this notion. Yet, unlike every version of this approach that I have found, my answer does not appeal to a *selective* sense of confirmation. This has been the rock upon which, arguably, every existing version of (C) has crashed. Therefore, there is no need to find an explication for the concept of rivalrous hypotheses that enables us to reduce the PR to a mere misunderstanding, because predictive confirmation can do the same job. Nonetheless, I must acknowledge inspiration from selective confirmation theorists like Goodman and Glymour. Additionally, predictive evidence is ‘selective’ in a different sense of the term, because the same evidence cannot (in normal contexts) predictively confirm both ‘All X are Y’ and ‘All X are  $\neg Y$ ’ because these hypotheses are associated with incompatible predictions: the first hypothesis is pragmatically associated with expecting that  $a$  satisfies Y, given the postulate that it satisfies X, whereas the second hypothesis is pragmatically associated with expecting that  $a$  satisfies  $\neg Y$ . Thus, predictive evidence selects among two contradictory predictions for an otherwise unknown individual  $a$ . Among their other merits, earlier versions of (C) were tantalizingly close to my answer.

## 5 Objections

One might wonder if my answer depends on idiosyncrasies of Bayesianism. In fact, my answer can be adapted to a variety of theories of evidence, provided that they can handle predictions in a satisfactory way. Hempel’s own system struggles in this regard (Hooker 1968) but my answer also works in Henry Kyburg’s system of “Evidential Probability” (Kyburg and Teng 2001). Unlike Bayesianism, this is a system of hypothesis acceptance and rejection, in which the fundamental core is a set of purely syntactic rules that govern the inference of hypotheses about relative frequencies. Firstly, according to Kyburg’s theory, universal generalisations of the form ‘All X are Y’ can be supported both by confirming that the frequency of Y among X’s is high or that the frequency of  $\neg X$  among  $\neg Y$  is high.<sup>19</sup> Secondly, singular predictions in Kyburg’s theory become more probable by accepting imprecise statistical generalisations given one’s evidence. In particular, if E confirms  $Y_a$  given  $(X_a \wedge B)$ , then E must confirm the hypothesis of a high relative frequency of Y in at least one reference class that we believe to contain  $a$ . Kyburg proposes various rules for determining which reference class(es) will be relevant, but for my definition we only assume that  $a$  is a member of the reference class of X’s. Therefore, it is only the relative

<sup>19</sup> In Kyburg’s view, the confirmation of universal generalisations involves a complex interplay of prediction, measurement, and convention: see Chapter 12 in *Uncertain Inference* (Kyburg and Teng 2001). However, the basic point I need is just that ‘All ravens are black’ can be confirmed by increasing the reliability of the inference of a high relative frequency of non-ravenness among non-black things.

frequency of Y in X's that is relevant. It is possible that evidence of non-black non-ravens might confirm *simpliciter* that 'All ravens are black', but only by supporting the statistical generalisation that the relative frequency of non-ravens among non-black things is high, rather than by confirming that the relative frequency of blackness among ravens is high. In Evidential Probability, as in Bayesianism, there can be divergences between confirmation *simpliciter* and the predictive component of predictive confirmation. My answer is not just an option for Bayesians.

In his discussion of the PR, Hempel considers and criticises an answer that is superficially similar to my own (Hempel 1945, 17–18). On that answer, the hypothesis 'All X are Y' has an implicit range of relevance, which is restricted to those things satisfying the expression X, and only instances in this range will confirm the hypothesis. I agree with Hempel that this is a mistake: it "involves a confusion of logical and practical considerations" (Hempel 1945, 18). The semantics of 'All ravens are black' has nothing particularly to do with ravens. However, that point is compatible with what I have said about predictive confirmation, where practical considerations have an independent and indispensable role. Therefore, it is unsurprising that the arguments that Hempel makes against the range of relevance answer do not apply to my answer. Firstly, he notes that scientists never make this range of relevance explicit, but on my answer the scope of the predictions associated with a universal generalisation (not the hypothesis itself) is suggested precisely by the choice of how one formulates the hypothesis: 'All X are Y' versus 'All  $\neg Y$  are  $\neg X$ '. Secondly, Hempel points out that there are commonplace logical operations (for instance, contraposition) which require that hypotheses of the forms 'All X are Y' and 'All  $\neg Y$  are  $\neg X$ ' have the same truth-conditions, but the range of relevance answer trades on distinct semantics for such hypotheses. In contrast, I have not denied that universal generalisations are contrapositable, but instead claimed that (in some circumstances) evidence for the reliability of one contrapositive's associated predictions might not be evidence for the reliability of the other contrapositive's associated predictions. Since this association is pragmatic, rather than semantic, it does not require a difference of truth-conditions.

My answer allows that 'All  $\neg Y$  are  $\neg X$ ' can be confirmed by evidence that instances of  $\neg Y$  are  $\neg X$ . Thus, 'All non-black things are non-ravens' can be confirmed, in the predictive sense, by a report of a non-black non-raven. Kyburg objects to confirmation theories with this feature, because scientists do not test hypotheses like 'All non-black things are non-ravens' by investigating the proportion of non-black things among non-ravens (Kyburg 1968, 309). Certainly, ornithologists do not test such hypotheses—but nor do they go around looking for black ravens to test 'All ravens are black'. Whether a rational scientist is interested in testing a hypothesis depends on a wide variety of factors, including the cost of testing, the probability of the hypothesis given the background evidence, its anticipated explanatory benefits, its expected technological utility, and so on. Therefore, it is possible that a particular hypothesis is testable, even though it would be silly to expend resources on testing it. 'All non-black things are non-ravens' could be such a hypothesis.

A different line of criticism could be made against the usefulness of predictive evidence. Why do I need to keep track of the reliability of making the predictions associated with a universal generalisation, given that the universal generalisation is well-confirmed? If I strongly believe that 'All ravens are black', then of course I also do not strongly believe that there are any non-black ravens. Surely universal generalisations can do all the necessary work; all my talk of 'the reliability of making the predictions associated with a universal generalisation' is redundant. I have two principal responses to this criticism: firstly, keeping track of the reliability of different predictive policies has a useful function of epistemic hygiene. In cases where 'All X are Y' is supported by

my evidence because I have good evidence that nothing satisfies X, I might forget that the reason I believe this hypothesis is not a predictively useful observed or hypothetical connection between X and Y given my total evidence, but simply because I had reasons to think that the hypothesis is vacuously satisfied. Keeping track of whether hypotheses are predictively confirmed, rather than merely confirmed, can help avoid such confusions. Douven (2008, 24) makes a similar point regarding the role of epistemic hygiene for the *acceptability* of conditionals. Secondly, recognising and retaining the reliability of different predictive policies helps prepare us for inferences after the loss of the universal generalisation: ‘All mammals do not lay eggs’ is no longer consistent with our evidence, but it is still a good rule-of-thumb, whereas ‘All Presidents of the United States of America are men’ will not be a good rule-of-thumb after there is a counterexample. For ideally rational agents, such advanced preparations for forgetfulness and rules of thumb are perhaps not important, but for flesh-and-blood humans, they are an inescapable part of our everyday reasoning.

My answer to the PR implies that confirmation is not a unitary concept, which might seem objectionable on grounds of complexity. However, there is precedent for taking confirmation to be ambiguous between multiple notions. For instance, Carnap (1962, xvi) distinguished between a variety of different sorts of confirmation, including both (1) whether a statement E increased the “firmness” of a hypothesis H given the relevant background information B and (2) whether H was “firm” on E and B. Another precedent of a non-unitary analysis of confirmation is that Joyce (2004, 144–145) uses a non-unitary analysis of confirmation to develop an intriguing answer to the Problem of Old Evidence for Bayesian epistemology. Simplicity can be sacrificed when there is a sufficient explicatory pay-off.

Finally, my answer to the PR is fundamentally empirical: the paradox is a product of an ambiguity in natural language. Yet I have only supported my answer through stylised facts associated with the PR and similar qualitative peculiarities concerning the analysis of universal generalisations. Therefore, one might reasonably worry that my answer is ad hoc. I have no novel evidence for my claims, but I can propose some experimental predictions. To begin, one would start by checking if each individual subject accepts the Scientific Laws Condition. Secondly, one would present the ravens hypothesis in a form that lacks the pragmatics that I have suggested are associated with universal generalisations in natural language, such as ‘Everything is a non-raven or black or both’, and check if the subjects understand the truth conditions of this sentence. (Given the doubtfully empirical status of ‘All ravens are black’, it might be preferable to use a hypothesis like ‘All panther mushrooms are poisonous’ and ‘Everything is a non-panther mushroom or poisonous or both’.) Finally, one could test to see if the PR survives the transformation: do people still find it counterintuitive that a non-black non-raven could be evidence for the hypothesis? My answer predicts that people would become comfortable with this possibility. A further prediction is that people who are troubled by a non-black non-raven confirming ‘All ravens are black’ will nonetheless generally be comfortable with the notion of such evidence confirming ‘All non-black things are non-ravens’, even though these generalisations are logically equivalent and have the same degrees of confirmation given the evidence. If I am correct, then a non-black non-raven does confirm ‘All non-black things are non-ravens’ relative to the implicit background information *and* confirms the reliability of making the predictions associated with it, and therefore I would expect that people generally do not find this paradoxical. I have no expertise in psychological testing, but it does seem that my explanation is testable and has some novel predictions. Still, I accept that it is sufficiently ad hoc to warrant significant scepticism, at least until we have tested its predictions beyond mere stylized facts and appeals to intuitions.

## 6 Conclusion

Our hesitance to say that reports of white shoes confirm that ‘All ravens are black’ is the product of an ambiguity. Once we disambiguate ‘confirms’ between confirmation *simpliciter* and predictive confirmation, we can happily say that (given certain background assumptions) we have confirmation *simpliciter* but not predictive confirmation for this hypothesis. Ordinary language often conflates these two types of evidence, yet formal explications of evidence are free to provide greater precision that can remove such paradoxes of ambiguity.

Critics of induction like Feyerabend (1968) and Popper (1974, 991) have used the PR to ridicule the notion of inductive reasoning. My answer implies that the paradox reveals no problems with induction at all. Abstracting from our ordinary inductive concepts can lead us astray *if* we fail to recognise what we are doing. This need for caution neither implies a problem for induction, nor a deep problem for abstract approaches to confirmation theory. I think that formally-orientated confirmation theory is perhaps the most successful research programme in all of philosophy, but we leave ourselves open to spurious paradoxes if we misunderstand the focus of this research. Pragmatics and formal analyses of confirmation theory can profitably travel together.

**Funding** The funding was provided by H2020 European Research Council (Grand No. 639276).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Carnap, R. (1962). *The logical foundations of probability*. London: Routledge & Kegan Paul.
- Couvalis, G. (1998). *The philosophy of science: Science and objectivity*. London: SAGE.
- Douven, I. (2008). The evidential support theory of conditionals. *Synthese*, 164(1), 19–44.
- Douven, I. (2017). How to account for the oddness of missing-link conditionals. *Synthese*, 194(5), 19–44.
- Douven, I., Elqyam, S., Singmann, H., & Wijnbergen-Huitink, J. V. (2018). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology*, 101, 50–81.
- Dretske, F. (1977). Laws of nature. *Philosophy of Science*, 44(2), 248–286.
- Feyerabend, P. K. (1968). A note on two ‘problems’ of induction. *The British Journal for the Philosophy of Science*, 19(3), 251–253.
- Fisch, M. (1984). Hempel’s ravens, the natural classification of hypotheses and the growth of knowledge. *Erkenntnis*, 21(1), 45–62.
- Fitelson, B., & Hawthorne, J. (2010). How bayesian confirmation theory handles the ravens paradox. In E. Eells & J. Fetzer (Eds.), *The place of probability in science* (pp. 247–275). Dordrecht: Springer.
- Franklin, J. (2001). Resurrecting logical probability. *Erkenntnis*, 55(2), 277–305.
- Glymour, C. (1980). *Theory and evidence*. Princeton: Princeton University Press.
- Goodman, N. (1954). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.
- Gower, B. (1997). *Scientific method: An historical and philosophical introduction*. London: Routledge.
- Hempel, C. G. (1945). Studies in the logic of confirmation I. *Mind*, 54(213), 1–26.
- Hooker, C. A. (1968). Goodman, ‘grue’ and Hempel. *Philosophy of Science*, 35(3), 232–247.
- Hosiasson-Lindenbaum, J. (1940). On confirmation. *The Journal of Symbolic Logic*, 5(4), 133–148.
- Jackson, F. (1987). *Conditionals*. Oxford: Blackwell.
- Joyce, J. M. (2004). Bayesianism. In A. R. Mele & P. Rawling (Eds.), *The Oxford handbook to rationality* (pp. 115–132). New York: Oxford University Press.

- Krzyżanowska, K., Wenmackers, S., & Douven, I. (2013). Inferential conditionals and evidentiality. *Journal of Logic, Language and Information*, 22(3), 315–334.
- Krzyżanowska, K., Wenmackers, S., & Douven, I. (2014). Rethinking Gibbard's riverboat argument. *Studia Logica*, 102(4), 771–792.
- Kyburg, H. E. (1968). *Philosophy of science: A formal approach*. New York: The Macmillan Company.
- Kyburg, H. E., & Teng, C. M. (2001). *Uncertain inference*. Cambridge: Cambridge University Press.
- Lipton, P. (2007). The ravens revisited. *Royal Institute of Philosophy Supplement*, 61, 75–95.
- Maher, P. (1999). Inductive logic and the ravens paradox. *Philosophy of Science*, 66(1), 50–70.
- Popper, K. (1974). Replies to my critics. In P. A. Schlipp (Ed.), *The philosophy of Karl Popper* (Vol. 2, pp. 961–1197). La Salle, IL: Open Court.
- Quine, W. V. O. (1970). Natural kinds. In N. Rescher (Ed.), *Essays in honour of Carl G. Hempel* (pp. 1–23). Dordrecht: D. Reidel.
- Ramsey, F. P. (1990). General propositions and causality. In D. H. Mellor (Ed.), *Philosophical papers* (pp. 145–163). Cambridge: Cambridge University Press.
- Rinard, S. (2014). A new bayesian solution to the paradox of the ravens. *Philosophy of Science*, 81(1), 81–100.
- Scheffler, I. (1968). *The anatomy of inquiry*. New York: Alfred A. Knopf.
- Schock, R. (1984). Ravens, grue, and material implication. *Dialectica*, 38(4), 347–350.
- Swinburne, R. G. (1971). The paradoxes of confirmation: A survey. *American Philosophical Quarterly*, 8(4), 318–330.
- Wright, G. H. (1966). The paradoxes of confirmation. In J. Hintikka & P. Suppes (Eds.), *Aspects of inductive logic* (pp. 208–219). Amsterdam: North-Holland Publishing Company.
- Zabell, S. L. (2004). Carnap and the logic of inductive inference. In D. M. Gabbay, S. Hartmann, & J. Woods (Eds.), *Handbook of the history of logic* (Vol. 10: Inductive logic, pp. 265–309). Oxford: Elsevier.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.