The Paradoxes of Confirmation

Jan Sprenger*

April 27, 2020

The **paradoxes of confirmation** are a group of generalizable examples that challenge the adequacy of specific formal accounts of when evidence confirms a theory. They show how these accounts classify intuitively spurious cases of confirmation as genuine evidential support, revealing problematic structural features of the involved confirmation criteria. In defending themselves against the paradoxes, confirmation theorists typically explain why our pre-theoretical intuitions about evidential support are mistaken, and they embrace the seemingly paradoxical conclusion (e.g., as Hempel did for the ravens paradox).

The paradoxes do not aim at accounts of confirmation that express whether hypothesis H is credible or acceptable in the light of observation E, such as Carnap's (1950) account of **confirmation as firmness**. Instead, they aim at accounts that capture whether E provides relevant **evidential support** for H—for example, by being predicted by H, by providing an instance of H, or by increasing the firmness of our degree of belief in H. In other words, they apply to probabilistic accounts such as Bayesian confirmation theory, but also to accounts based on deductive relations in first-order logic, such as hypothetico-deductive confirmation and Hempel's satisfaction criterion.

The paradoxes of confirmation are a group of three examples: (1) the paradox of the ravens, also known as Hempel's paradox; (2) Goodman's

*Contact information: Center for Logic, Language and Cognition (LLC), Department of Philosophy and Educational Science, Università degli Studi di Torino, Via Sant'Ottavio 20, 10124 Torino, Italy. Email: jan.sprenger@unito.it. Webpage: www.laeuferpaar.de.

new riddle of induction, also known as the "grue" paradox; (3) the tacking paradoxes, or more specifically, the problem of irrelevant conjunctions and disjunctions. Not all of them affect each account of confirmation. The **paradox of the ravens** arises most forcefully for naïve theories of confirmation by instantial relevance, such as: observing a black raven confirms the hypothesis that all ravens are black. Adding plausible additional assumptions, it then follows that this hypothesis is also confirmed by observations such as a white shoe. Also **Goodman's new riddle of induction** aims in the first place at instance-based accounts of confirmation: he argues a particular observation (e.g., a green emerald) confirms mutually incompatible hypotheses that make wildly divergent claims about the future (e.g., that emeralds examined in the future will be green, blue, red, etc.). Finally, the **tacking paradoxes** show that on a hypothetico-deductivist account, the confirmation relation is maintained when an irrelevant conjunct (e.g., "the moon consists of green cheese") is tacked to the confirmed hypothesis. Table 1 gives an overview of which account of confirmation and evidential support is affected by which paradox of confirmation.

	Paradoxes of Confirmation		
Accounts of Evidential Support	Paradox of the Ravens	New Riddle of Induction	Tacking Paradoxes
Naïve Instantial Relevance (Nicod)	yes	yes	no
Refined Instantial Relevance (Hempel)	yes	yes	no
Hypothetico-Deductive (H-D) Model	(no)	(no)	yes
Increase in Firmness (Bayes)	yes	yes	yes

Table 1: An overview of the most prominent accounts of confirmation as evidential support and their relation to the three paradoxes of confirmation.

Thus, a paradox of confirmation has to be discussed relative to a specific account. Our discussion will also reveal some surprising findings. While the Bayesian account of confirmation as increase in firmness is affected by all three paradoxes, we will see that it is also good at developing responses, due to the inherent flexibility of the probabilistic framework.

The chapter is structured as follows. Section 1 discusses the paradox of the ravens by explaining its historical genesis and formal analysis. Section 2 does the same for Goodman's new riddle of induction. Section 3 shows how the tacking paradoxes challenge H-D confirmation and how they are mitigated on a Bayesian account. Section 4 draws general morals and concludes.

1 The Paradox of the Ravens

Natural laws, and hypotheses about natural kinds, are often formulated in the form of universal conditionals, such as "all planets move in elliptical orbits", "all ravens are black" or "all tigers are predators". According to a longstanding tradition in philosophy of science, such hypotheses are confirmed by their instances (Nicod 1925/61; Hempel 1945a, 1965b; Glymour 1980): the hypothesis "all F's are G's" is confirmed by the observation of a F that is also a G ($Fa.Ga$). This suggests the following condition first mentioned by Jean Nicod:

Nicod Condition (Confirmation by Instances) Universal conditionals such as $\mathrm{H} = \forall x \colon (Fx \rightarrow Gx)$ are confirmed by their instances, that is, propositions such as $\mathrm{E} = Fa \wedge Ga$.

At the same time, formal theories of confirmation should respect some elementary logical principles. For example, if two hypotheses are logically equivalent, they should be equally confirmed or undermined by a given observation E. This brings us to the

Equivalence Condition If observation E confirms (undermines) hypothesis H, then it also confirms (undermines) any hypothesis H′ that is logically equivalent to H.

Hempel (1945a,b) observed that combining the Equivalence and the Nicod Condition leads to paradoxical results. Take the hypothesis that no non-black object is a raven: $\mathrm{H}' = \forall x : \neg Bx \rightarrow \neg Rx$. A white shoe is an instance of that hypothesis. Thus, by the Nicod Condition, observing a white shoe ($\mathrm{E}' = \neg Ba \wedge \neg Ra$) confirms H′. By the Equivalence Condition, H′ is equivalent to $\mathrm{H} = \forall x : Rx \rightarrow Bx$ so that E′ also confirms the hypothesis that all ravens are black. But why on Earth should the observation of a white shoe matter for our attitude toward the color of ravens?

Ravens Intuition Observations of a white shoe or other non-black non-ravens do *not* confirm the hypothesis that all ravens are black.

Hence, we have three individually plausible, but incompatible claims—the Nicod Condition, the Equivalence Condition and the Ravens Intuition—at least one of which has to be discarded. Since this paradox of the ravens was first formulated by Carl G. Hempel in his essays "Studies in the Logic of Confirmation" (I+II, cited above and reprinted in Hempel 1965a), it is also known as **Hempel's paradox**.

Hempel suggests to give up the Ravens Intuition and to embrace the paradoxical conclusion. Assume that we observe a grey bird that resembles a raven. This observation puts the raven hypothesis at risk: the bird may be a non-black raven and falsify our hypothesis. However, by conducting a genetic analysis we learn that the bird is no raven, but a kind of crow. Here, it sounds correct to say that the results of the genetic analysis support the raven hypothesis—it was at risk of being falsified and has survived a test (=the genetic analysis, see also Popper 1959/2002).

This way of telling the story explains why observations of the form $\neg Ra \wedge \neg Ba$ can confirm the raven hypothesis $\mathrm{H} = \forall x : Rx \rightarrow Bx$. But why did we have a different intuition in the first place? Hempel thinks that this is due to an ambiguity in the way the paradox is set up. In the above crow/raven case, we did not yet know whether the newly observed bird was a raven or a crow. Therefore its investigation has confirmatory potential. By contrast, in the white shoe example, we know that *the object before us is no raven*:

> [...] this has the consequence that the outcome of the [...] test becomes entirely irrelevant for the confirmation of the hypothesis and thus can yield no new evidence for us. (Hempel 1945a, 19)

That is, the observation of a white shoe should better be described as the observation of a non-black object ($\mathrm{E}' = \neg Ba$) relative to the background knowledge that the object is not a raven ($\mathrm{K}' = \neg Ra$). The Nicod Condition is not satisfied for E′ relative to H′ and K′ and so the paradox vanishes. Hempel's analysis explains in particular why "indoor ornithology" (e.g., looking for white shoes) does not yield evidential support for the raven hypothesis.

The presence of relevant background knowledge can thus make a huge difference to whether or not confirmation obtains. According to Hempel, confirmation should therefore not be modelled as a two-place relation between hypothesis H and evidence E, but as a three-place relation between H, E and background knowledge K. Following this line of thought, Hempel replaces the Nicod Condition by a more refined and formalized account of instantial relevance:

Satisfaction criterion (Hempel) A piece of evidence E **directly Hempel-confirms** a hypothesis H relative to background knowledge K if and only if E and K jointly entail the development of H to the domain of E. In other words, $E \wedge K \models H_{|dom(E)}$.

The development of a hypothesis to a set of objects $\{a, b\}$ is the set of predictions the hypothesis makes if restricted to these entities. Thus, the development of $\forall x : Fx$ to the domain of $E = Fa \wedge Fb \wedge Ga \wedge \neg Ha$ is $Fa \wedge Fb$. As Goodman (1955/83, 69) puts it:

> a hypothesis is genuinely confirmed only by a [observation] statement that is an instance of it in the special sense of entailing not the hypothesis itself but its relativization or restriction to the class of entities mentioned by that statement

Unfortunately, as pointed out by Fitelson and Hawthorne (2011), Hempel's constructive proposal does not match his analysis of the paradox. Since the satisfaction criterion is *monotonic* with regard to background knowledge, the confirmation relation remains intact when background knowledge is added. In particular, when we do know beforehand that a is no raven and observe a to be non-black ($E = \neg Ba$, $K = \neg Ra$), it will still be the case that

$$E \wedge K = \neg Ra \wedge \neg Ba \models (Ra \rightarrow Ba) = H_{|dom(E)}.$$

Thus, even when we know beforehand that E is irrelevant for K, H ends up being confirmed on Hempel's account. While Hempel spots correctly

Distribution of Objects	Scenario 1 (S_1)	Scenario 2 (S_2)
Black ravens	100	1,000
Non-black ravens	0	1
Other objects	1,000,000	1,000,000

Table 2: I. J. Good's (1967) counterexample to the Nicod Condition shows how observing an instance of an universal conditional ("all ravens are black") can lower its probability.

that the paradoxical conclusion of the raven example can be embraced by relegating the paradoxical aspect to implicit background knowledge, his own theory of confirmation does not implement that insight.

For the Bayesian account of confirmation as increase in firmness, things are more complicated. According to that theory, confirmation or evidential support corresponds to an increase in subjective probability or degree of belief (see also the chapter on positive relevance by Peter Achinstein):

Confirmation as Increase in Firmness For hypothesis H, observation E and background knowledge K, E **confirms/supports** H relative to K if and only if E raises the probability of H conditional on K: $p(H \mid E, K) > p(H \mid K)$, and vice versa for disconfirmation.

The Bayesian account of confirmation as probability-raising tells us something more than Hempel's original analysis: it enables us to spot that instantial relevance may be a bad guide to confirmation. Indeed, the probability of a universal conditional can also be *lowered* by observing its instances. I. J. Good (1967) proposed a simple example to this effect, reproduced in Table 2. We compare two scenarios, S_1 and S_2, where objects are either black ravens, non-black ravens, or something else (e.g., grey crows).

Suppose that we can rule out any other scenario on the basis of our background knowledge. Then the hypothesis $H = \forall x \colon (Rx \rightarrow Bx)$ that all ravens are black is true in S_1 and false in S_2. Moreover, since there are more black ravens in S_2 than in S_1, observing a black raven raises the probability that S_2 is the case:

$$p(Ra \wedge Ba \mid S_1) = \frac{100}{1{,}000{,}100} < \frac{1{,}000}{1{,}001{,}001} = p(Ra \wedge Ba \mid S_2).$$

Due to the symmetry of probabilistic relevance and our assumption $S_1 \leftrightarrow \neg S_2$, we can infer $p(S_1 \mid Ra \wedge Ba) < p(S_1)$, and equivalently, $p(H \mid Ra \wedge Ba) < p(H)$. Thus, observing a black raven may also *disconfirm* the raven hypothesis. The explication of confirmation as increase in firmness corrects our pre-theoretic intuitions about the validity of the Nicod Condition (compare Sprenger and Hartmann 2019, 51–53).

However, the Bayesian needs to explain why we *actually* find the observation of a black raven better evidence for the raven hypothesis than a non-black non-raven. In this spirit, Fitelson and Hawthorne (2011) show in their Theorem 2 (op. cit.) that *typically*—that is, for plausible assumptions on the base rate of ravens and black objects—observing a black raven provides better evidential support than a non-black non-raven. In other words, $Ra \wedge Ba$ raises the probability of the raven hypothesis H more than $\neg Ba \wedge \neg Ra$ does: $p(H \mid Ra \wedge Ba) > p(H \mid \neg Ra \wedge \neg Ba)$. This means that $Ra \wedge Ba$ confirms H to a higher degree than $\neg Ba \wedge \neg Ra$ for all evidential support measures that depend only on the prior and posterior probability of H. This includes the most important measures in the literature (Crupi 2015; Sprenger and Hartmann 2019, chapter 1; see also the chapter on confirmation measures in this volume). Ultimately, Fitelson and Hawthorne's result reveals why a black raven is more important evidence for the raven hypothesis than a white shoe.

While Fitelson and Hawthorne arguably resolve the **comparative version of the ravens paradox**, many authors have been aiming at a stronger result, that is, to show that the observation of a white shoe (or in general, a non-black non-raven) provides almost zero confirmation for the hypothesis that all ravens are black (e.g., Horwich 1982; Earman 1992; Howson and Urbach 1993). Such arguments for resolving the **quantitative version of the ravens paradox** rest, however, on disputable assumptions such as $p(Ba \mid H, K) \approx p(Ba \mid \neg H, K)$, that is, the truth of the ravens hypothesis barely affects the probability that a randomly sampled object is black (for discussion, see Vranas 2004). For these reasons, the Bayesian response to the paradox is best summarized by (1) rejecting the Nicodian intuition that instances always confirm universal generalizations and (2) showing that under typical circumstances, a black raven confirms the raven hypothesis more than a white shoe. More discussion of Bayesian and non-Bayesian approaches to the paradox of the ravens can be found in Maher 1999, Huber 2007 and Sprenger 2010.

2 The New Riddle of Induction

Goodman devised his "new riddle of induction" in the third chapter of "Fact, Fiction and Forecast" (Goodman 1955/83) as a challenge to Hempel's (1945a; 1945b) account of confirmation, the satisfaction criterion. However, it can be framed as a general problem: only lawlike statements should be confirmed by their instances and formal theories of confirmation usually can't tell lawlike from accidental generalizations. While there is considerable discussion about how Goodman's paradox should be understood (e.g., Jackson 1975; Okasha 2007; Fitelson 2008), we adopt a confirmation-theoretic reading where the paradox shows how hypotheses with incompatible predictions are confirmed by the same evidence.

The presentation below follows Sprenger 2016, pp. 190–191, and Sprenger and Hartmann 2019, pp. 53–54. Consider the following case of a standard inductive inference:

$$\begin{array}{l} \text{Observation at } t = t_1\text{: emerald } e_1 \text{ is green.} \\ \text{Observation at } t = t_2\text{: emerald } e_2 \text{ is green.} \\ \qquad\qquad\qquad\qquad \vdots \\ \text{Observation at } t = t_n\text{: emerald } e_n \text{ is green.} \\ \hline \text{Conclusion: All emeralds are green.} \end{array}$$

The conclusion follows by the premise by enumerative induction (i.e., the straight rule of induction): the observation of the emeralds confirms the general hypothesis that all emeralds are green. The inference seems intuitively valid and Nicod's and Hempel's confirmation criteria agree.

Goodman now declares an object to be *grue* either (1) if it is green and has been observed up to time $t_{\text{now}} = t_n$, or (2) it is blue and is observed for the first time after t_{now}. Notably, no object is required to change color to count as grue. Consider now the following inductive inference:

$$
\begin{array}{l}
\text{Observation at } t = t_1\text{: emerald } e_1 \text{ is grue.} \\
\text{Observation at } t = t_2\text{: emerald } e_2 \text{ is grue.} \\
\qquad\qquad\qquad\qquad \vdots \\
\text{Observation at } t = t_n\text{: emerald } e_n \text{ is grue.} \\
\hline
\text{Conclusion: All emeralds are grue.}
\end{array}
$$

The inference that all emeralds are grue looks fishy, but formally, it is based on the very same rule (i.e., enumerative induction) as the previous inference that all emeralds are green. More specifically, the observation of a green emerald prior to t_{now} is an instance of the "green" as well as the "grue" hypothesis and thus Hempel-confirms both hypotheses. It is therefore not clear which inferences and expectations about the future are licensed by the observations. While Hempel's paradox showed that confirmation by instances does not exclude evidence which we find intuitively irrelevant, Goodman's paradox shows that the same observations confirm incompatible hypotheses and cast doubt of the ability of a purely formal account of confirmation to support rational expectations about the future.

A natural reaction to the paradox is to deny the use of the predicate "grue" in inductive inferences due to their explicit use of temporal restrictions. Goodman (1955/83, 79–80) responds that such a move would be arbitary: we can redefine the standard predicates "green" and "blue" in terms of "grue" and its conjugate predicate "bleen": an object is green if (1) it is grue and observed prior to t_{now} or (2) it is bleen and observed after t_{now}. Any preference for the "natural" predicates and the "natural" inductive inference is a relative, not an absolute matter, and conditional on the choice of a specific language. For purely formal accounts of confirmation such as Nicod's confirmation by instances and Hempel's satisfaction criterion, it is therefore difficult to avoid the conclusion that observing green emeralds confirms the hypothesis that all emeralds are grue.

Goodman suggests to drive a wedge between "green" and "grue" by restricting the confirmation relation to predicates with successful prediction history. Instead, one could also accept that evidence can confirm incompatible hypotheses. For example, Einstein's work on the photoelectric effect confirmed the hypothesis that light is composed of discrete quanta (photons) rather than a contionuous wave. Thereby Einstein's discovery also confirmed different and mutually incompatible versions of quantum theory (e.g., relativistic and non-relativistic versions).

The Bayesian's answer to Goodman's paradox follows this latter line: she admits that both hypotheses are supported by the observed evidence, but denies that they are supported *equally*. Evidential support typically depends on some function of prior and posterior probability, and the "green" hypothesis is for evident reasons more plausible than the "grue" hypothesis. Moreover, many measures of evidential support validate the **Matthew effect** (Festa 2012; Festa and Cevolani 2017): evidential support is, ceteris paribus, higher for hypotheses with high prior plausibility, favoring the "green" over the "grue" hypothesis. Fitelson (2008) offers a different Bayesian solution.

The choice of priors, however, is external to Bayesian inference: they are motivated by theoretical principles, past observations and scientific judgment. The rules of Bayesian inference state how we should amalgamate prior degrees of belief with observed evidence, but they do not state which prior degrees of belief are reasonable. Thus, Goodman's paradox highlights the need for inductive assumptions in inductive inference and shows that Bayesian Confirmation Theory is not an inductive *perpetuum mobile* (see also Norton 2018).

3 The Tacking Paradoxes

So far, we have been silent on one of the principal and most venerable models of confirmation in science: the **hypothetico-deductive (H-D) model**. It is at the basis of a lot of scientific practice that sees a theory as confirmed if its predictions obtain. For example, the prediction that light would be bent by massive bodies like the sun, and Eddington's verification of this prediction during the 1919 eclipse was widely seen as a strong confirmation of Einstein's General Theory of Relativity. The idea that successful and risky predictions are essential to our assessment of a scientific hypothesis was also put forward famously by Karl R. Popper (1959/2002, 1963). Another intuition that supports this approach to confirmation is the prediction-accommodation asymmetry: we typically prefer hypotheses with a good predictive track record over those which have been fitted *ad hoc* to the data (Whewell 1847; Worrall 1989; Hitchcock and Sober 2004).

The H-D model regards a hypothesis as confirmed if predictions have been derived deductively from the hypothesis, and the hypothesis was necessary to make these predictions:

Hypothetico-Deductive (H-D) Confirmation Observation E H-D-confirms hypothesis H relative to background knowledge K if and only if

1. $H \wedge K$ is consistent;
2. $H \wedge K$ entails E ($H \wedge K \models E$);
3. K alone does not entail E.

The H-D approach to confirmation avoids the paradox of the ravens because the raven hypothesis $H = \forall x : Rx \rightarrow Bx$ does not make any predictions about whether a particular object a is a black raven or a non-black non-raven. The only thing H entails is that a cannot be a non-black raven. We can also reconstruct Hempel's analysis of the paradox: relative to the background knowledge $K_1 = Ra$, the observation $E_1 = Ba$ H-D-confirms the raven hypothesis. So does the observation $E_2 = \neg Ra$ relative to background knowledge $K_1 = \neg Ba$—for example, the genetic analysis of a grey bird whom we cannot classify by the eye as crow or raven. In other words, black ravens, grey crows and even white shoes can all confirm the raven hypothesis as long as they the observation constitutes a genuine test of the hypothesis. So the paradox of the ravens is handled in a satisfactory way by the H-D account.

Similarly, Goodman's paradox does not arise because the H-D account *embraces* the conclusion that contradictory hypotheses can be confirmed by the same evidence. Remember that the paradoxical aspect of Goodman's paradox emerged from the fact that hypotheses with incompatible predictions are confirmed by the evidence. This is, however, only worrying for "inductivist" or "projectivist" accounts of confirmation such as instantial relevance and probability-raising. By contrast, the rationale behind H-D confirmation is in the first place to assess the past performance of hypotheses in experimental tests (see also Popper 1959/2002, ch. 10).

A more serious challenge for the H-D account is given by the so-called **tacking paradoxes**. Suppose H denotes the General Theory of Relativity, E denotes Eddington's observation during the 1919 solar eclipse that star light is bent by the sun, and X a nonsensical hypothesis such as "The moon consists of green cheese". H is hypothetico-deductively confirmed by observation E relative to background knowledge K, and so is $H \wedge X$ because this hypothesis also entails E, due to the monotonicity of logical entailment. Thus, the *conjunction* of General Theory of Relativity and the hypothesis that the moon consists of green cheese has been confirmed. This sounds completely absurd since confirmation is transmitted "for free" to the irrelevant, and in fact, nonsensical conjunct which has never been tested empirically.

The scheme can be generalized easily to other examples. The bottom line is that if E H-D-confirms H, then E also H-D-confirms any $H \wedge X$ that is consistent with H and K. When we tack an *irrelevant conjunct* to a confirmed hypothesis, H-D-confirmation is preserved. This is highly unsatisfactory. An analogous (though less discussed) problem for H-D confirmation is the confirmation of a hypothesis H by evidence that is logically weaker than the deductively implied prediction E, such as the *disjunction* of E with an irrelevant observation O, that is, $E \vee O$. In both cases, it seems that H-D confirmation misses out on a satisfactory account of evidential relevance (Moretti 2006).

There have been several solution proposals trying to endow H-D confirmation with an account of evidential relevance. Early proposals by Horwich (1982) and Grimes (1990) must be regarded as having failed (e.g., Gemes 1998). Two more promising accounts have been developed in later work. The first is the account of **relevant conclusions** by Gerhard Schurz (1991), based on the idea that any predicate that relevantly occurs in a logical inference must contribute to the inference and cannot be replaced arbitarily by a different predicate *salva veritate*. This account classifies an entailment such as $\forall x : Fx \models Fa \vee Ga$ as irrelevant since the entailment would remain valid if G were replaced by any other predicate. An analogous definition can be given for irrelevant premises. The definition of H-D confirmation is then amended by the requirement that the entailment $\mathrm{H} \wedge \mathrm{K} \models \mathrm{E}$ contains neither irrelevant premises nor irrelevant conclusions.

The second proposal is based on the concept of **content parts**. The basic idea, introduced by Ken Gemes (1993), is that an entailment $\mathrm{H} \models \mathrm{E}$ is relevant if and only if every *relevant model* of E (i.e., a model that assigns truth values only to those atomic formulae that affect the truth value of E) can be extended to a relevant model of H. In such a case, we call E a content part of H and amend the second clause of H-D confirmation by demanding that E be a content part of $\mathrm{H} \wedge \mathrm{K}$. This modification takes care of the problem of tacking irrelevant disjuncts to E; for tacking irrelevant conjuncts to H, complementary solutions have been suggested by Gemes (1993) and Sprenger (2013). More discussion of H-D confirmation and possible fixes for the tacking paradoxes is found in Sprenger 2011.

For an account of confirmation based on instantial relevance (Nicod, Hempel), the tacking paradoxes do not arise since these accounts do not validate what Hempel calls the **Converse Consequence Condition** (i.e., if E confirms H, it also confirms logically stronger hypotheses). The Bayesian, however, has to respond to the tacking paradoxes since confirmation as increase in firmness subsumes H-D confirmation as a special case. If H entails E conditional on background knowledge K, it will be the case that $p(\mathrm{E}|\mathrm{H},\mathrm{K}) = 1$. By assumption, then also $p(\mathrm{E}|\mathrm{K}) < 1$ since otherwise, K would already have entailed E, contradicting the assumptions of H-D confirmation. Thus, $p(\mathrm{E}|\mathrm{H},\mathrm{K}) > p(\mathrm{E}|\mathrm{K})$ and by Bayes' Theorem, we infer $p(\mathrm{H}|\mathrm{E},\mathrm{K}) > p(\mathrm{H}|\mathrm{K})$ and E confirms H relative to K. The Bayesian replies, as usual, that the problem can be *mitigated* from a comparative point of view: If E raises the probability of hypothesis H, and E is intuitively relevant for H, the degree of evidential support is higher than for the hypothesis $\mathrm{H} \wedge \mathrm{X}$ where X denotes, as before, an irrelevant conjunct. Obviously, such a claim is sensitive to the used measure of evidential support (see the chapter by Peter Brössel in this volume). For example, for the ratio measure $r(\mathrm{H}, \mathrm{E}, \mathrm{K}) = p(\mathrm{H}|\mathrm{E}, \mathrm{K})/p(\mathrm{H}|\mathrm{K})$ we can derive

$$\begin{aligned} r(\mathrm{H} \wedge \mathrm{X}, \mathrm{E}, \mathrm{K}) &= p(\mathrm{H} \wedge \mathrm{X} \,|\, \mathrm{E}, \mathrm{K})/p(\mathrm{H} \wedge \mathrm{X} | \mathrm{K}) \\ &= p(\mathrm{E} \,|\, \mathrm{H} \wedge \mathrm{X} \wedge \mathrm{K})/p(\mathrm{E}|\mathrm{K}) = 1/p(\mathrm{E}|\mathrm{K}) = p(\mathrm{E}|\mathrm{H}, \mathrm{K})/\mathrm{p}(\mathrm{E}|\mathrm{K}) = \mathrm{r}(\mathrm{H}, \mathrm{E}, \mathrm{K}), \end{aligned}$$

showing that the conjunction $\mathrm{H} \wedge \mathrm{X}$ is supported to the same degree as the original hypothesis H. The other principal measures of evidential support like the difference or the likelihood ratio measure fare better in this respect: whenever $p(\mathrm{E} \,|\, \mathrm{H} \wedge \mathrm{X} \wedge \mathrm{K}) = p(\mathrm{E} \,|\, \mathrm{H} \wedge \mathrm{K})$, they reach the conclusion that $\mathrm{H} \wedge \mathrm{X}$ is confirmed less than H (Hawthorne and Fitelson 2004, revised Theorem 2). This result on the tacking paradoxes includes H-D confirmation as a special case since $p(\mathrm{E} \,|\, \mathrm{H} \wedge \mathrm{X} \wedge \mathrm{K}) = p(\mathrm{E}|\mathrm{H} \wedge \mathrm{K}) = 1$ whenever H entails E. Bayesian Confirmation Theory acknowledges the tacking paradoxes, but demonstrates at the same time how they can be mitigated.

4 Conclusion

The paradoxes of confirmation show how the absence of a theory of evidential relevance challenges purely formal accounts of confirmation: a syntactic criterion detects evidential support, but the case looks spurious to our pre-theoretical intuitions. "All ravens are black" is confirmed by observing a white shoe, "all emeralds are grue" by observing a green emerald, and observations of the planetary motions confirm Kepler's laws together with the hypothesis that the moon consists of green cheese.

The paradoxes consist of two groups: Goodman's and Hempel's paradoxes are particularly problematic for accounts of confirmation based on instantial relevance (Nicod, Hempel) whereas the tacking paradoxes attack accounts based on the concept of successful prediction, such as the H-D account (see again Table 1). Immunization strategies have no easy life: while Hempel's account of confirmation fails to conform with his own analysis of the raven paradox, the H-D account can block the tacking paradoxes only at the price of considerable technical complications.

The Bayesian account of confirmation as increase in firmness is in general more permissive than those accounts (e.g., it includes H-D confirmation as a special case), so it has to respond to all three paradoxes. On the other hand, the Bayesian has an additional strategy at her disposal, unavailable to purely qualitative ("all-or-nothing") accounts of confirmation: she admits that intuitively irrelevant evidence may raise the probability of the hypothesis, but that it does so to a lesser extent than the relevant contrast class (e.g., black raven/white shoe). This **comparative resolution** is applicable to all three paradoxes—but it falls short of showing that the degree of support provided in the problematic case is close to zero. It is much more demanding to provide such a **quantitative resolution**. While the best treatment of the paradoxes remains an open research question, the existing results show that the paradoxes do not trivalize confirmation theory and that the latter remains a fruitful branch of philosophy of science and formal epistemology.

Acknowledgements

This chapter rewrites excerpts from Sprenger 2010, Sprenger 2011 and the first chapter of Sprenger and Hartmann 2019 and merges them with novel material. I would like to thank Bengt Autzen and Peter Brössel for their valuable feedback, and the editors of this volume, Clayton Littlejohn and Maria Lasonen-Aarnio, for their invitation to contribute. Research on this article was supported through Starting Investigator Grant "Making Scientific Inferences More Objective" (grant No. 640638) by the European Research Council.

References

Carnap, Rudolf (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Crupi, Vincenzo (2015). Confirmation. In Ed Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/archives/sum2018/entries/confirmation/.

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

Festa, Roberto (2012). *"For Unto Every One that Hath Shall be Given"*: Matthew Properties for Incremental Confirmation. *Synthese 184*, 89–100.

Festa, Roberto and Gustavo Cevolani (2017). Unfolding the grammar of Bayesian confirmation: likelihood and anti-likelihood principles. *Philosophy of Science 84*, 56–81.

Fitelson, Branden (2008). Goodman's 'New Riddle'. *Journal of Philosophical Logic 37*, 613–643.

Fitelson, Branden and James Hawthorne (2011). How Bayesian Confirmation Theory Handles the Paradox of the Ravens. In James H. Fetzer and Ellery Eells (eds.), *The Place of Probability in Science*, pp. 247–275. New York: Springer.

Gemes, Ken (1993). Hypothetico-Deductivism, Content and the Natural Axiomatisation of Theories. *Philosophy of Science 60*, 477–487.

Gemes, Ken (1998). Hypothetico-Deductivism: The Current State of Play; the Criterion of Empirical Significance: Endgame. *Erkenntnis 49*, 1–20.

Glymour, Clark (1980). *Theory and Evidence*. Princeton, N.J.: Princeton University Press.

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

Goodman, Nelson (1955/83). *Fact, Fiction and Forecast* (4th ed.). Cambridge, Mass.: Harvard University Press.

Grimes, Thomas R. (1990). Truth, content, and the hypothetico-deductive method. *Philosophy of Science 57*, 514–522.

Hawthorne, James and Branden Fitelson (2004). Re-Solving Irrelevant Conjunction with Probabilistic Independence. *Philosophy of Science 71*, 505–514.

Hempel, Carl G. (1945a). Studies in the Logic of Confirmation I. *Mind 54*, 1–26.

Hempel, Carl G. (1945b). Studies in the Logic of Confirmation II. *Mind 54*, 97–121.

Hempel, Carl G. (1965a). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.

Hempel, Carl G. (1965b). Science and Human Values. In *Aspects of Scientific Explanation*, pp. 81–96. New York: Free Press.

Hitchcock, Christopher and Elliott Sober (2004). Prediction versus Accommodation and the Risk of Overfitting. *British Journal for the Philosophy of Science 55*, 1–34.

Horwich, Paul (1982). *Probability and Evidence*. Cambridge: Cambridge University Press.

Howson, Colin and Peter Urbach (1993). *Scientific Reasoning: The Bayesian Approach* (2nd ed.). La Salle, Ill.: Open Court.

Huber, Franz (2007). Confirmation and Induction. In Fieser J. and Dowden B. (eds.), *Internet Encyclopedia of Philosophy*. Retrieved from https://www.iep.utm.edu/conf-ind/ on April 27, 2020.

Jackson, Frank (1975). Grue. *Journal of Philosophy 72*, 113–131.

Maher, Patrick (1999). Inductive Logic and the Ravens Paradox. *Philosophy of Science 66*, 50–70.

Moretti, Luca (2006). The tacking by disjunction paradox: Bayesianism versus hypothetico-deductivism. *Erkenntnis 64*, 115–138.

Nicod, Jean (1925/61). *Le problème logique de l'induction*. Paris: Presses Universitaires de France. Originally published in 1925 (Paris: Alcan).

Norton, John D. (2018). A Demonstration of the Incompleteness of Calculi of Inductive Inference. *British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axx004.

Okasha, Samir (2007). What Does Goodman's 'Grue' Problem Really Show? *Philosophical Papers 36*, 483–502.

Popper, Karl R. (1959/2002). *The Logic of Scientific Discovery*. London: Routledge. Reprint of the revised English 1959 edition. Originally published in German in 1934 as "Logik der Forschung".

Popper, Karl R. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge.

Schurz, Gerhard (1991). Relevant Deduction. *Erkenntnis 35*, 391–437.

Sprenger, Jan (2010). Hempel and the Paradoxes of Confirmation. In Dov Gabbay, Stephan Hartmann, and John Woods (eds.), *Handbook of the History of Logic*, Volume 10, pp. 235–263. Amsterdam: North-Holland.

Sprenger, Jan (2011). Hypothetico-Deductive Confirmation. *Philosophy Compass 6*, 497–508.

Sprenger, Jan (2013). A Synthesis of Hempelian and Hypothetico-Deductive Confirmation. *Erkenntnis 78*, 727–738.

Sprenger, Jan (2016). Confirmation and Induction. In Paul W. Humphreys (ed.), *Handbook of Philosophy of Probability*, pp. 185–209. Oxford: Oxford University Press.

Sprenger, Jan and Stephan Hartmann (2019). *Bayesian Philosophy of Science*. Oxford: Oxford University Press.

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

Whewell, William (1847). *Philosophy of the Inductive Sciences, Founded Upon Their History*. London: Parker.

Worrall, John (1989). Structural Realism: The Best of Both Worlds? *Dialectica 43*, 99–124.