Hempel and Confirmation Theory

Jan Sprenger*

June 15, 2020

Carl Gustav Hempel (1905–1997) was one of the primary exponents of logical empiricism. As a student and member of the *Gesellschaft für empirische Philosophie* in Berlin, alongside Reichenbach and Grelling, he witnessed the emergence of logical empiricism as a philosophical program. From the mid-1930s onwards, his contributions shaped its development, too. Hempel studied primarily in Göttingen and Berlin, but in 1929/30, he also spent a semester in Vienna studying with Carnap and participated in the activities of the Vienna Circle. Both societies joined forces for organizing scientific events, and founded the journal *Erkenntnis* in 1930, where many seminal papers of logical empiricism were published, with Carnap and Reichenbach as editors.

While the work of the Berlin philosophers is congenial to the project of the Vienna Circle, there are important differences, too. Neither Hempel nor his mentor Reichenbach identified "scientific philosophy" with the project of cleansing science of meaningless statements (e.g., Carnap 1930). Rather, Hempel extensively used a method that Carnap would apply in later works on probability and confirmation (Carnap 1950, 1952): *explication*, that is, the replacement of a vague and imprecise pre-theoretical concept (e.g., "confirmation") by a fruitful and precise concept (e.g., a formal confirmation criterion). Relying on the method of explication, Hempel developed adequacy conditions on a qualitative concept of confirmation (Hempel 1943, 1945a,b), a probabilistic measure of degree of

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

confirmation (Hempel and Oppenheim 1945), and most famously, the D-N model for explanation by means of natural laws (Hempel and Oppenheim 1948). Much of contemporary philosophy of science, including formally oriented literature (e.g., Sprenger and Hartmann 2019), is closer to Hempel's approach of explicating central concepts in ordinary scientific reasoning than it is to Carnap's reconstructive approach in the *Aufbau* (Carnap 1928/1998), or his later work on logical foundations of inductive inference (e.g., Carnap 1950). That said, Hempel and Carnap share the conviction that we must analyze the relationship between theory and evidence not only in terms of verifying the observable consequences of a theory, but also, and specifically, in terms of the *inductive* consequences of a given body of evidence for the assessment of a theory or hypothesis. More than Carnap, Hempel also worked on scientific reasoning in a broader context, especially in later years. Specifically, he engaged with Rudner's 1953 provocative thesis that the scientist's acceptance of a hypothesis always involves value judgments, and commented extensively on the most influential works from the next generation: Thomas S. Kuhn's "Structure" (Kuhn 1962) and Paul Feyerabend's "Against Method" (Feyerabend 1975).

This chapter gives an overview of Hempel's work on confirmation and induction. Section 1 explains Hempel's take on the problem of induction and his probabilistic explication of degree of confirmation. We then proceed, in Section 2, to Hempel's explication of the classificatory or qualitative concept of confirmation, and the Satisfaction Criterion in particular. Section 3 presents the famous paradox of the ravens, Hempel's analysis and its impact on later work. Finally we briefly review Goodman's "new riddle on induction" that takes issue with Hempel's confirmation criteria, and Hempel's later work on values in inductive inference. The final Section 4 concludes.

1 The Modern Problem of Induction

Students of philosophy all learn about Hume's classical problem of induction: how to justify beliefs and actions that are based on empirical, logically inconclusive evidence. According to Hempel, however, a second problem of induction is at least of equal importance: specifying the rules for a valid inference from empirical evidence (the premises) to a theoretical hypothesis (the conclusion), or in other words, finding a **logic of inductive inference**. Such a logic would try to mirror the success of deductive logic for ampliative inferences, and secure the objectivity of inductive inferences in science. Specifically, it would replace the subjective appraisal of a theory by objective, verifiable standards for confirmation (Hempel and Oppenheim 1945, 98–99) and contribute to the central aims of empiricist philosophy: to understand and to model the progress of science, the replacement of old by new theories, and the testability of abstract hypotheses by empirical observations.

Hempel (1965b, 30–34) is quick to point out that inductive inference cannot consist in an indiscriminate collection of facts, followed by their systematization and inductive generalizations. In fact, like Popper, Hempel stresses that the scientific process must be guided by tentative hypotheses which we later evaluate on the basis of empirical evidence. And like Popper, Hempel insists that a logic of scientific reasoning cannot cover this essentially creative and non-regulated process of inventing hypotheses. Neither can it prescribe the decision to accept or reject a hypothesis on the basis of evidence: for Hempel, this decision is entangled with pragmatic values (Hempel 1965c, 1983). "Rules of inductive inference will have to be conceived, not as canons of discovery, but as criteria of validation for proposed inductive arguments" (Hempel 1965b, 34): they do not *generate* a hypothesis from a given body of evidence, but *presuppose* that a hypothesis (or varying competing hypotheses) has been put forward independently, and evaluate that hypothesis against the available evidence. It is here, in the comparison of theoretical sentences that express a hypothesis, with observation reports expressing the evidence, that logical tools can make an important contribution.

In this context, Hempel stresses the importance of Carnap's (1947) Requirement of Total Evidence (RTE): the inductive support in favor of a hypothesis H should be calculated with respect to the total available evidence E. However, for Hempel the RTE is no rule of inductive inference (like, e.g., "observed deductive consequences of a theory confirm it"), but a rule that governs the *rational application* of inductive inferences (Hempel 1965b, 43). As such, it is especially salient in contexts where we would like to base a decision on accepting or rejecting a theory on the quantitative degree of support $dc(\mathrm{H}, \mathrm{E})$ in favor of H.

Hempel and Oppenheim (1945) propose to measure inductive support based on the method of *maximum likelihood* that R. A. Fisher (1935/74) introduced into statistics few years before. The idea is to find, for any evidence E, the "optimum distribution" Δ_E over the probability space, that is, the distribution that assigns maximal probability to E. Then H is assigned, as degree of confirmation with respect to E, the probability of H under Δ_E, that is, $dc(\mathrm{H}, \mathrm{E}) = p_{\Delta_\mathrm{E}}(H)$ (my notation, J.S.).

This procedure is, of course, well-known from maximum likelihood estimation: as a point estimate of an unknown parameter θ, one chooses the value $\hat{\theta}$ that assigns maximal probability to the observed data. And like maximum likelihood, Hempel and Oppenheim's method need not yield unique results, as the authors notice. For example, when the elements of the probability space are unrelated propositions of predicate logic, with $\mathrm{H} = Fa$ and $\mathrm{E} = Gb$, then any degree of confirmation $0 \leq dc(\mathrm{H}, \mathrm{E}) \leq 1$ will be admissible. Hempel and Oppenheim's criterion will often just determine interval bounds for the inductive support of E for H. This lack of uniqueness is not necessarily a vice, however: also Carnap (1952) defended in later work the multiplicity of inductive methods and moreover, Hempel and Oppenheim show that the thus defined degree of confirmation obeys several intuitive principles, such as:

- If H is a logical consequence of E (and E is consistent), then $dc(\mathrm{H}, \mathrm{E}) = 1$.
- For any optimum distribution Δ, $dc(\mathrm{H}, \mathrm{E}) + dc(\neg\mathrm{H}, \mathrm{E}) = 1$.

The function $dc(\mathrm{H}, \mathrm{E})$ acts in many respects like a probability function, and it can be connected to various plausible constraints on degree of confirmation. It differs from Carnapian confirmation functions (Carnap 1950) in various ways, though: it is not meant to explicate the pre-theoretical concept of "probability", it is strongly inspired by statistical reasoning, and it does (unlike Carnap's confirmation functions) not depend on the partitioning of the logical space.

2 Qualitative Confirmation Criteria

While Carnap quantified degree of confirmation without addressing the question of when a piece of evidence confirms a theory at all, Hempel thought that **qualitative adequacy conditions** were a necessary prolegomenon for a quantitative, probabilistic account of confirmation (Hempel 1945a, 30–33). Such adequacy conditions are supposed to capture the core elements of the concept of confirmation, and to constrain the quantitative analysis of confirmation in a successive stage. The first condition Hempel proposes is the

Entailment Condition (EnC) If hypothesis H logically follows from the observation report E, then E confirms H.

For example, if the hypothesis reads "there are black ravens" then, the observation of a single black raven *proves* it and *a fortiori*, confirms it: logical implication is the strongest possible form of evidential support.

Then, in an inductive *logic*, confirmation should extend to the logical consequences of what is already confirmed. For instance, if we have evidence for Newton's law of gravitation, it must also be evidence for Kepler's laws, since the latter are a special case of the former. In other words, Hempel suggests the

Consequence Condition (CC) If an observation report E confirms every member of a set of sentences $\mathcal{S}$, then it confirms every logical consequence of $\mathcal{S}$, too (e.g. every sentence H for which $\mathcal{S} \models$ H).

The Consequence Condition also implies the

Special Consequence Condition (SCC) If an observation report E confirms a hypothesis H, then it confirms every logical consequence of H, too.

The SCC clashes, however, with an important intuition about the link between prediction and confirmation: instead of testing theories directly, we often verify their observational consequences. For instance, the General Theory of Relativity was first tested by checking its predictions for the bending of light by massive bodies. This predictivist approach to confirmation motivates the

Converse Consequence Condition (CCC): If an observation report E confirms a hypothesis H, then it confirms every hypothesis H′ that logically implies H (i.e. $H' \models H$).

Accepting both SCC and CCC, however, would trivialize the concept of confirmation. (By EnC, E confirms E; by CCC, E confirms $E \wedge H$ for any H; by SCC, E then confirms H—even when no actual link between E and H exists.) Faced with this choice, Hempel opts for SCC and dismisses CCC. Mainly because CCC extends the confirmation relation too generously: it allows for the confirmation of mutually incompatible hypotheses (if E confirms H, then E confirms both $H \wedge X$ and $H \wedge \neg X$), and because of its vulnerability to the tacking paradoxes for hypothetico-deductive confirmation (if E confirms H, then E confirms $H \wedge X$ for any X, see e.g., Gemes 1998; Sprenger 2011b). In fact, for an inductive *logic* it is a strange feature that incompatible conclusions follow from the same set of premises. In line with this reasoning, Hempel adopts the

Consistency Condition (CnC) If an observation report E confirms the hypotheses H and H′, then H′ must be logically consistent with H (i.e., there are models of H′ that are also models of H).

The three cornerstornes of Hempel's qualitative adequacy criteria are thus the Entailment Condition, the (Special) Consequence Condition and the Consistency Condition. Hempel then combines these formal criteria with a substantial confirmation criterion. Of course, logical entailment is usually too strong as a necessary criterion for confirmation: no finite set of observations will ever imply a universal statement of the form "all Fs are Gs". However, the hypothesis should agree with the evidence in the domain of the evidence. Specifically, Hempel suggests that if an observation report says something about a set of the singular terms (e.g., S_E =

$\{a, b, c\}$), the evidence should provide a model of the restriction or development of the hypothesis to S_E (a precise definition is given in Hempel 1943). For instance, if $E = Fa \wedge Ga \wedge Fb$, then the development of $H = \forall x : Fx \rightarrow Gx$ to $S_E = \{a, b\}$ is $H_{|dom(E)} = (Fa \rightarrow Ga) \wedge (Fb \rightarrow Gb)$. This brings us to the

Satisfaction Criterion A piece of evidence E *directly Hempel-confirms* a hypothesis H if and only if E provides a model of the restriction of H to the domain of E. In other words, $E \models H_{|dom(E)}$, where $H_{|dom(E)}$ denotes the restriction of H to the singular terms that occur relevantly in E.

This criterion can be generalized as follows: anything that follows classically from a set of directly confirmed hypotheses counts as confirmed, in agreement with Hempel's Consequence Condition.

Hempel-Confirmation A piece of evidence E *Hempel-confirms* a hypothesis H if and only if H is entailed by a set of sentences Γ so that for all sentences $\phi \in \Gamma$, ϕ is directly Hempel-confirmed by E.

It is easy to see that Hempel's account satisfies the three above criteria. It also improves upon several shortcomings of both the naïve account of confirmation by instances, and the hypothetico-deductive account. In fact, it also stands at the core of Glymour's (1980) account of bootstrap confirmation, and it can be connected to hypothetico-deductive confirmation, too (Sprenger 2013). We will now move to the raven paradox as an important test case for formal theories of confirmation.

3 The Ravens' Paradox

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 lions are carnivores". According to a tradition in philosophy of science that goes back to Jean Nicod (1925/61), hypotheses of the form "all *F*'s are *G*'s" are confirmed by their *instances*, that is, observations of an F that are also Gs (e.g., E = $Fa \wedge Ga$). This suggests the following condition:

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, as we have seen in the last section, theories of confirmation should respect certain logical principles. For example, if two hypotheses H and H′ are logically equivalent, they should be equally confirmed by an observation E: inductive support should not depend on the chosen formulation of a hypothesis. This brings us to the

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

In fact, the Equivalence Condition is not only highly plausible; it also follows directly from Hempel's other criteria: if H is equivalent to H′, H also implies H′. Thus, if E confirms H, E also confirms H′ by SCC. However, Hempel (1945a,b) observed that combining the Equivalence and the Nicod Condition runs counter to established confirmatory intuitions. Take the hypothesis that no non-black object is a raven: $\mathrm{H}' = \forall x : \neg Bx \rightarrow \neg Rx$. A white swan is an instance of that hypothesis. Thus, by the Nicod Condition, observing a white swan ($\mathrm{E}' = \neg Ba \wedge \neg Ra$) 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 obviously, observing a white swan should not affect our attitude toward the color of ravens.

Ravens Intuition Observations of a white swan 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 Hempel in his essays "Studies in the Logic of Confirmation" I+II (=Hempel 1945a,b, reprinted in Hempel

1965a), it is also known as **Hempel's paradox**. Even before, Hempel (1937, 221–222) proposed a similar counterexample for measuring the degree of confirmation of universal conditionals—see also Hosiasson-Lindenbaum 1940.

Facing this trilemma, Hempel dismisses the Ravens Intuition and embraces the paradoxical conclusion: observing a white swan confirms the hypothesis that all ravens are black. To motivate that resolution, assume that we observe a grey bird that resembles a raven. This 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). This Popperian line of response is also worked out by various papers in the 1950s and 1960s (e.g., Watkins 1957; Agassi 1958; Good 1966; Hempel 1967).

Hempel's analysis explains why white swans, or more generally, observations of the form $\neg Ra \wedge \neg Ba$, can confirm the raven hypothesis. But why did we have a different intuition in the first place? Hempel traces this back to an ambiguity in the paradox. In the crow/raven case, we did not yet know whether the newly observed bird was a raven or a crow. Therefore its investigation has confirmatory (and falsificatory) potential. By contrast, in the white swan 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 swan should better be described as the observation of a non-black object ($E' = \neg Ba$) relative to the background knowledge that the object is not a raven ($K' = \neg Ra$) and such cases, that do not put the hypothesis at risk, should *not* count as confirming instances (cf. Popper 1959/2002). $E = \neg Ba \wedge \neg Ra$ and $K = \emptyset$, by contrast confirms the raven hypothesis. Accounting for background knowledge explains in particular why "indoor ornithology" cannot yield support for the raven hypothesis. Following this road we have to rewrite the above confirmation criteria accounting for the role of background knowledge. For the Satisfaction Criterion, this can be done straightforwardly (i.e., the main condition becomes $E \wedge K \models H_{dom(E)}$).

Fitelson and Hawthorne (2011) pointed out that Hempel's Satisfaction Criterion does not square well with his analysis of the paradox. Since the Satisfaction Criterion is *monotonic* with regard to background knowledge, adding background knowledge cannot invalidate inductive support. In particular, even when we know that a is no raven ($E = \neg Ba$, $K = \neg Ra$), it will still be the case that

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

Thus, even for irrelevant evidence E, the raven hypothesis H is confirmed on Hempel's account. While Hempel spots correctly 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.

The raven paradox also anticipates Nelson Goodman's **new riddle of induction**. Goodman set up this problem in the third chapter of *Fact, Fiction and Forecast* (Goodman 1955/83) as a challenge to the Satisfaction Criterion. If we observe only green emeralds up to time point $t = t_0$, this should, in any plausible logic of inductive inference, support the hypothesis that all emeralds are green. Hempel's Satisfaction Criterion agrees in fact if we formalize the hypothesis as the universal conditional $H : \forall x : Ex \rightarrow Gx$.

However, consider now the predicate "grue" which applies to an emerald e either (1) if e is green and has been observed up to time t_0, or (2) if e is blue and is observed for the first time after t_0. We can then redescribe our past observations as "emerald e_1 is grue", "emerald e_2 is grue", and so on. These observations support, according to the very same rules of inductive inference—the Satisfaction Criterion in particular—, the hypothesis H′ that all emeralds are grue. This conclusion violates the Consistency Condition: both H and H′ are confirmed by the same observations. Moreover, they are not only incompatible with each other, but also disagree on each single prediction for $t > t_0$. These are highly undesirable consequences. While Hempel's paradox shows that observing instances of a hypothesis is no reliable guide to inductive inference, Goodman's new riddle demonstrates that for any universal generalization H that is confirmed according to the Satisfaction Criterion, we can construe a rival hypothesis H' such that the same observations confirm H' although H and H' make completely incompatible predictions. This casts doubt of the ability of a purely formal, syntactic account of confirmation to support rational expectations about the future.

In a postscript to the "Studies in the Logic of Confirmation" that appeared as part of Hempel 1965a, Hempel admits that the Consistency Condition may be too strong as an adequacy criterion for an inductive logic. And in fact, it is no coincidence that in later works, Hempel did not work any more on formal confirmation criteria. Specifically, in his article "Science and Human Values", Hempel (1965c) took up arguments by Rudner (1953) and others that non-cognitive, properly ethical values influence the decisions and inferences made by scientists. While Hempel stays faithful to his earlier views that ethical values do not have a logical relationship to theory and evidence, and therefore do not affect judgments of inductive support, he stresses that the *acceptance* or rejection of a hypothesis always carries a risk of error—the famous "inductive risk". Weighing these errors is not a purely logical process and needs to be done on the basis of definite utilities and losses assigned to correct and erroneous decisions. This later work by Hempel has been very influential recently, for example in Heather Douglas's work (Douglas 2000, 2009).

Finally, Hempel's late work shows a certain degree of scepticism toward the original logical empiricist project: the application of scientific theories for purposes of explanation and prediction depends on auxiliary assumptions, so-called "provisos"—for example, the absence of factors that could interfere with the forces postulated by the theory (Hempel 1988). In the light of this additional complexity, the task of formulating purely syntactic accounts of confirmation, explanation and inductive inference becomes increasingly difficult.

4 Conclusion

Carl Gustav Hempel has been an ingenious researcher with manifold contributions to the development of 20th-century philosophy of science. The field of confirmation and induction is no exception. Some of his contributions in that area are nowadays mainly of historical interest: for example, the Satisfaction Criterion has mainly been superseded by hypothetico-deductive and Bayesian accounts of confirmation. Nonetheless, as one of the first systematic attempts to develop formal criteria for inductive inference, Hempel's work inspired important follow-up research, such as Goodman's "new riddle of induction", or Glymour's (1980) bootstrap confirmation. Hempel's work on the paradox of the ravens and the analysis he provides, by contrast, are seminal up to today and continue to generate numerous original research articles. From a methodological point of view, his insistence on developing adequacy criteria before moving to a quantitative analysis of confirmation has proven to be an extremely helpful strategy, and it is followed also in various parts of Bayesian philosophy of science (e.g., Fitelson 2001; Sprenger and Hartmann 2019). All in all, Hempel's contributions to the problem of induction and confirmation theory may not be as deep and detailed as Carnap's, but they equal them in terms of originality and interest, and they exceed them in terms of breadth of perspective.

Suggestion for further reading

Many of Hempel's original articles on confirmation and induction are collected in two volumes edited by James Fetzer (2000, 2001). Apart from ordering Hempel's papers thematically and selecting his most important contributions, these books also provide extensive introductions to Hempel's thought and compare Hempel's work to other philosophers of logical empiricism. Crupi (2020) provides a useful overview of confirmation theory, including Hempel's contributions; Sprenger (2011a) surveys later work on Hempel's paradox.

Acknowledgements

This chapter rewrites excerpts from Sprenger 2011a and Sprenger forthcoming and merges them with novel material. I would like to thank the editors of this volume, Christoph Limbeck-Lilienau and Thomas Uebel, for their invitation to contribute, and their valuable feedback. Research on this article was supported through Starting Investigator Grant "Making Scientific Inferences More Objective" (grant No. 640638) by the European Research Council, and PRIN grant "From Models to Decisions" by the Italian Ministry for Universities and Research.

References

Agassi, Joseph (1958). Corroboration versus induction. *British Journal for the Philosophy of Science 9*, 311.

Carnap, Rudolf (1928/1998). *Der logische Aufbau der Welt*. Felix Meiner. Original edition published in 1928 by Weltkreis Verlag, Berlin.

Carnap, Rudolf (1930). Die alte und die neue Logik. *Erkenntnis 1*, 12–26.

Carnap, Rudolf (1947). On the Application of Inductive Logic. *Philosophy and Phenomenological Research 8*, 133–148.

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

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

Crupi, Vincenzo (2020). Confirmation. In Ed Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/archives/spr2020/entries/confirmation/.

Douglas, Heather (2000). Inductive Risk and Values in Science. *Philosophy of Science 67*, 559–579.

Douglas, Heather (2009). *Science, Policy, and the Value-Free Ideal*. Pittsburgh: Pittsburgh University Press.

Fetzer, James (2001). *The Philosophy of Carl G. Hempel*. Studies in Science, Explanation and Rationality. Oxford: Oxford University Press.

Fetzer, James H. (2000). *Science, Explanation, and Rationality: Aspects of the Philosophy of Carl G. Hempel*. Oxford University Press.

Feyerabend, Paul (1975). *Against Method*. London: Verso.

Fisher, R. A. (1935/74). *The Design of Experiments*. New York: Hafner Press. Reprint of the ninth edition from 1971. Originally published in 1935 (Edinburgh: Oliver & Boyd).

Fitelson, Branden (2001). *Studies in Bayesian Confirmation Theory*. PhD thesis, University of Wisconsin–Madison.

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 (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. (1966). 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. Originally published in 1955.

Hempel, Carl G. (1937). Le problème de la vérité. *Theoria 3*, 206–244.

Hempel, Carl G. (1943). A Purely Syntactical Definition of Confirmation. *Journal of Symbolic Logic*.

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). Recent Problems of Induction. In R. G. Colodny (ed.), *Mind and Cosmos*, pp. 112–134. Pittsburgh: Pittsburgh University Press. Reprinted in Hempel, Carl G. and Fetzer, James: "The Philosophy of Carl G. Hempel", pp. 29–48.

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

Hempel, Carl G. (1967). The white shoe: No red herring. *British Journal for the Philosophy of Science 18*, 239–240.

Hempel, Carl G. (1983). Valuation and objectivity in science. In R. Cohen and L. Laudan (eds.), *Physics, Philosophy, and Psychoanalysis*, pp. 73–100. D. Reidel.

Hempel, Carl G. (1988). Provisos: A Problem Concerning the Inferential Function of Scientific Theories. *Erkenntnis 28*, 147–164.

Hempel, Carl G. and Paul Oppenheim (1945). A Definition of 'Degree of Confirmation'. *Philosophy of Science 12*, 98–115.

Hempel, Carl G. and Paul Oppenheim (1948). Studies in the Logic of Explanation. *Philosophy of Science 15*, 135–175.

Hosiasson-Lindenbaum, Janina (1940). On confirmation. *The Journal of Symbolic Logic 5*, 133–148.

Kuhn, Thomas S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

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

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".

Rudner, Richard (1953). The Scientist *qua* Scientist Makes Value Judgments. *Philosophy of Science 20*, 1–6.

Sprenger, Jan (2011a). 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 (2011b). 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 (forthcoming). The paradoxes of confirmation. In Clayton Littlejohn and Maria Lasonen-Aarnio (eds.), *Routledge Handbook of the Philosophy of Evidence*. London: Routledge.

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

Watkins, J. W. N. (1957). Between analytic and empirical. *Philosophy 32*, 112–131.