14
CONFIRMATION

Alan Hájek and James M. Joyce

Introduction, motivation, central concepts

Introduction

Confirmation theory is intended to codify the evidential bearing of observations on hypotheses, characterizing relations of inductive *support* and *counter-support* in full generality. The central task is to understand what it means to say that datum *E* confirms or supports a hypothesis *H* when *E* does not logically entail *H*.

While there were important investigations into confirmation theory by Bacon, Whewell, Mill, and Duhem, the modern study of confirmation was pioneered by Hempel (1945) and Carnap (1962). Given its importance to the philosophy of science and to epistemology, it is surprising that philosophy had to wait so long for well-developed theories of confirmation. This may have been due to a general skepticism about the possibility of inductive support stemming from Hume's *problem of induction*. Hume famously questioned our entitlement to infer things about the future from our experience of the past, and his skeptical arguments can be generalized to cover all non-deductive inferences. More recently, Popper's deductivist philosophy of science has been equally unfriendly to confirmation theory.

Yet the denial of non-deductive confirmation relations is tantamount to skepticism. Without such relations, you have no right to infer the existence of an external world from your perceptions, nor the existence of other minds from the existence of your own, nor anything about your past from your apparent memories of it. Whatever its philosophical credentials, confirmation theory is deeply rooted in common sense, and rational decision and science would be impossible without the relations that are its subject matter.

Concepts of confirmation

Confirmation theorists countenance two relations of confirmation, characterized by the following schemata:

Absolute: *H* is highly supported given evidence *E*.
Incremental: *E* increases the evidential support for *H*.

Both notions assume a background of total evidence. "*E* is absolute evidence for *H*" means that given *E*, the total evidence for *H* lies above some salient threshold. "*E* incrementally confirms H" means that adding E to the background data increases the total evidence for H. It is important to recognize that E can be incremental evidence for H without being absolute evidence for H, and conversely. For example, testing positive for AIDS provides incremental evidence that you have AIDS, but it may not provide absolute evidence: it may be more likely that the test has produced a false positive than it is that you have AIDS.

The ordinary notion of *confirmation* seems to involve both incremental and absolute elements, neither fully accounting on its own for our speech or our practices. Even so, we will focus largely on incremental confirmation, taking the notion of absolute confirmation as understood.

Theories of confirmation

Qualitative confirmation

Hempel thought that the development of the "logic" of confirmation should proceed in stages: qualitative, comparative, and quantitative. He encountered problems at the first stage. In keeping with his logical empiricism, he sought to characterize confirmation in largely deductive terms. His 1945 article presents the following conditions as *prima facie* plausible:

1 *Entailment condition*: If E implies H, then E confirms H.
2 *Special consequence condition*: If E confirms H, and H implies H', then E also confirms H'.
3 *Special consistency condition*: If E confirms H, and H is incompatible with H', then E does not confirm H'.
4 *Converse consequence condition*: If E confirms H, and if H is implied by H', then E also confirms H'.

But, as Hempel recognized, any relation satisfying 1–4 will hold between *every* pair of propositions, clearly an unacceptable result. (Actually, 1 and 4 jointly suffice for the unacceptable result, as Moretti (2003) observes.) Hoping to preserve as much of 1–4 as possible within a unitary account of confirmation, Hempel restricted 4 to cases where H is obtained from H' by instantiation, while maintaining 1–3.

Carnap (1962, new Preface) argues that Hempel conflated incremental and absolute confirmation. In any case, while 1–3 are plausible for absolute confirmation, 4 is not – e.g., the fact that H is well supported given E does not imply that the conjunction of H with some highly unlikely proposition is also well supported given E. The situation regarding incremental confirmation is more nuanced. 3, a squarely *absolutist* principle, clearly fails. 1, 2 and 4, which mix absolute and incrementalist intuitions, hold only in special cases, albeit important ones: 1 fails when H already has probability 1, but otherwise holds. 2 breaks down when E increases the evidence for H while more strongly decreasing the evidence for $H'\&\neg H$, but it holds when E either supports or is irrelevant to $H'\&{\sim}H$. 4 fails when E increases the evidence for $H'\&{\sim}H$

while decreasing the evidence for H' by a smaller amount, but it holds when $H' = H$ & X for X an "irrelevant conjunct" that is not evidentially germane to either H or E (Fitelson 2002).

Instance confirmation and the ravens paradox

Hempel also endorses a famous condition that is *prima facie* plausible for incremental confirmation, but completely implausible for absolute confirmation:

> *Nicod's condition*: All universal generalizations of the form "All *F*s are *G*" are confirmed by all statements of the form "*a* is both *F* and *G*."

For example, it seems plausible that the report of a particular black raven incrementally confirms the generalization "All ravens are black," but implausible that the report absolutely confirms the generalization.

A special case of 2 (and of 4), and compelling in its own right, is the *equivalence condition*:

> If H is logically equivalent to H', and E confirms H, then E also confirms H'.

Nicod's condition and the equivalence condition yield Hempel's notorious *ravens paradox*. Since "All ravens are black" is equivalent to "All non-black things are non-ravens," Nicod's condition apparently entails that the latter generalization is confirmed by the report of the observation of any non-black non-raven, e.g., a white shoe. But by the equivalence condition, "All ravens are black" is likewise confirmed by any such report. This seems paradoxical: white shoes seem to have no evidential bearing whatsoever on ornithological hypotheses.

Hempel embraces the paradox, arguing that our intuitions recoil only because we know that there are far more non-black things than ravens. Confirmation relations, on Hempel's view, should presuppose no such background knowledge. Good (1967) replies that a confirmation theory that ignores knowledge is of little interest to science. But once we make confirmation a three-place relation, with background knowledge as the third relatum, Nicod's criterion plainly fails – see sub-section "Probability theory and probabilistic measures of support."

Quine (1969) argues that Nicod's criterion is false insofar as it quantifies over *all* predicates *F* and *G*. He insists that confirmation relations must be restricted to *natural kind* predicates, those whose instances are objectively similar to each other. While "raven" and "black" are plausibly natural kind predicates, "non-raven" and "non-black" are not (their miscellaneous instances including electrons and quasars). Alternatively, one might regard Quine as casting doubt on the Equivalence condition: while "All ravens are black" is apt for confirmation, "All non-black things are non-ravens" is not.

As we shall see, such extreme remedies seem like overkill on probabilistic approaches to confirmation. But first we must consider their best-known rival.

H-D confirmation

Hypothetico-deductivism is perhaps the most familiar and historically influential confirmation theory. Its more sophisticated forms, e.g., Ayer (1936), are motivated by the thought that a hypothesis is confirmed by data it entails, but are tempered by the recognition that entailments between hypotheses and data are almost always mediated by background knowledge.

> *H-D confirmation*: E incrementally confirms H iff there are true "auxiliary hypotheses" $A_1, A_2, \ldots, A_n$ such that (a) $A_1 \,\&\, A_2 \,\&\, \ldots \,\&\, A_n$ does not entail E, while (b) $H \,\&\, A_1 \,\&\, A_2 \,\&\, \ldots \,\&\, A_n$ entails E but not $\sim E$.

Unfortunately, as Duhem (1905) already recognized, auxiliary hypotheses that figure in confirmation relations are, like the hypothesis under test, fallible conjectures based on inconclusive evidence. This led Quine (1951) to insist that confirmation is *holistic*, i.e., that evidence never confirms or disconfirms any hypothesis in isolation. *H-D* confirmation is thus restricted to "total theories" with enough content to entail observations on their own. While such total theories are confirmed by their empirical consequences, their individual hypotheses are not. This has the unpalatable result that there is no principled way to differentially distribute praise or blame over hypotheses.

Another serious challenge to hypothetico-deductivism, in either its holistic or atomistic form, is the *underdetermination of theory by evidence*. Moreover, the model does not address *statistical hypotheses*, since these have no empirical consequences (e.g., any pattern of "heads" and "tails" is compatible with a coin's being fair). As these problems illustrate, H-D confirmation is not sufficiently nuanced to isolate the evidential relationships we care about. For those we need to invoke probabilities.

Probabilistic theories of confirmation

Probability theory and probabilistic measures of support

Probabilistic theories of confirmation assume that claims of confirmation and disconfirmation must be evaluated relative to some probability function (or set of such functions), which encodes all the background information relevant in a context of inquiry. A probability function P is an assignment of real numbers to elements of some set S of propositions, closed under negation and countable disjunction and conjunction, obeying the following axioms (for all $A, B \in S$):

1 $P(A) \geq 0$.
2 $P(T) = 1$ for all logical truths T.
3 $P(A1 \vee A2 \vee \ldots) = P(A1) + P(A2)$ when all the Ai and Aj are contraries.
4 The probability of A conditional on B is given by

$$P(A \mid B) = \frac{P(A \,\&\, B)}{P(B)}, \text{ provided } P(B) > 0.$$

If P encapsulates all of an agent's opinions and background knowledge, then $P(H)$ reflects the total evidence for H based on her prior knowledge alone, while $P(H|E)$ reflects the evidence for H when E (and nothing else) is added to that knowledge. In contrast, $P(E)$ and $P(E|H)$ convey information about E's predictability: $P(E)$ reflects E's predictability based on what is known; $P(E|H)$ reflects its predictability when H (and nothing else) is added to this knowledge. Conditional probabilities can thus be used to reflect either the epistemic status of a hypothesis in light of potential data or the predictive power of the hypothesis with respect to that data.

Probabilistic theories represent increases in evidential support using relations of probabilistic relevance and independence. At the qualitative level, the idea is that confirming evidence raises the probability of a hypothesis, disconfirming evidence lowers it, and irrelevant evidence leaves it unchanged:

> *Probabilistic theory of incremental evidence* (*qualitative*): Relative to probability function P,
> - E incrementally confirms H iff $P(H|E) > P(H)$.
> - E incrementally disconfirms H iff $P(H|E) < P(H)$.
> - E is evidentially irrelevant to H iff $P(H|E) = P(H)$.

This simple theory has some appealing consequences:

- Evidence for a hypothesis is always evidence against its negation.
- Most H-D confirmation is probabilistic confirmation since $P(H|E)$ exceeds $P(H)$ when H entails E unless $P(H)$ or $P(E)$ equal 0 or 1.
- E increases the evidence for H iff H increases E's predictability.

The probabilistic approach also provides a useful framework for understanding the effect of background information on confirmation. To see how, let's revisit the raven paradox. On a probabilistic picture, instance confirmation is straightforward:

> *Probabilistic IC*: $\forall x(Fx \supset Gx)$ is incrementally confirmed by any learning experience in which (a) one of its *logical* instances $\sim Fa \vee Ga$ becomes certain, (b) there was some positive prior probability that a is both F and $\sim G$, and (c) nothing else of relevance is learned.

Let H be $\forall x(Fx \supset Gx)$. Intuitively, $\sim Fa \vee Ga$ confirms H by ruling out a "live" counterexample in which $P(Fa \,\&\, \sim Ga) > 0$. Because it relies on a logically weaker notion of an instance, probabilistic-IC has significant advantages over Nicod's condition. Here are two:

- Given (a)−(c), $\sim Fa \vee Ga$ always confirms both $\forall x(Fx \supset Gx)$ and $\forall x(\sim Gx \supset \sim Fx)$.
- $\sim Fa \vee Ga$ increases the evidential support for H only if there is a non-zero probability that a is both F and $\sim G$.

(c) deserves special attention since much of the raven paradox's *paradoxicality* can be traced directly to it. Probabilistic IC implies that learning that some object is either a non-raven or black, *and nothing more*, always raises the probability of *H*. But experience often delivers *additional* information, whose effect on *H*'s probability can vary greatly depending on the information encoded in *P*. Suppose we are sampling birds, at random and with replacement, from a fixed population of 1,000, and consider the following states of prior knowledge:

(i) Either 950 birds are ravens but only 949 of these are black, or 10 birds are ravens and all are black (Good 1967).
(ii) 998 birds are black ravens. At least one of the other two is white, but it is unknown whether either is a raven.
(iii) 900 birds are black ravens. All the others are white, but it is unknown whether any are ravens.
(iv) There are 990 ravens, 980 already known to be black. Of the 20 remaining birds either 10 are black ravens and 10 are white doves, or all are ravens, each equally likely to be white or black.
(v) There are at most 50 ravens. Ten ravens have been found to be black. The rest of the population is heterogeneous with respect to color.

Suppose that probabilities equal the corresponding proportions. In (i), observing a black raven *lowers* *H*'s probability, whereas observing a non-black non-raven raises *H*'s probability. In (ii), a non-black non-raven *raises* *H*'s probability. A black raven also raises *H*'s probability, but less so. In (iii), a black raven does not alter *H*'s probability at all, but something known only to be a non-black non-raven increases it. In (iv), a black raven raises *H*'s probability slightly. Something known only to be neither black nor a raven lowers *H*'s probability. But a white non-raven raises $P(H)$ to 1! Case (v) is most like the one in which we find ourselves. Observing either a black raven or a non-black non-raven raises *H*'s probability, but since there are vastly more non-black things than ravens, the increase is much greater for the first observation than for the second.

In all these cases, information beyond that found in $\sim Ra \lor Ba$ has a significant effect on confirmation relations. Depending on the background information, such extra information can alter the probability of the hypothesis in almost any way. Moreover, this information can be about white shoes, red herrings, or anything else. For instance, if we know that all ravens are black iff white shoes exist, then observing a white shoe verifies the hypothesis. This does not, however, conflict with the intuition that "All ravens are black" can only be confirmed by evidence about ravens. Information about non-ravens can, given the right background knowledge, also be evidence about ravens. The raven paradox seems paradoxical only when we fail to appreciate this point.

The dependence of prior probability on background information also offers some relief from the Duhem–Quine problem. Suppose that the conjunction of *H* and auxiliary hypothesis *A* entails $\sim E$, and that *E* is observed. Depending on *P*, *E* may:

- decrease *H*'s probability but not *A*'s (or vice versa);
- increase *H*'s probability but decrease *A*'s (or vice versa);
- decrease both *H*'s and *A*'s probability.

The question of which of these occurs depends on the prior probabilities of the four conjunctions of *H*/~*H* and *A*/~A, and on the predictability of *E* when these combinations are assumed. See Earman (1992) for discussion.

Probabilistic approaches also facilitate discussion of confirmation in comparative and quantitative terms through *Bayes's theorem*:

$$\frac{P(H|E)}{P(H)} = \frac{P(E|H)}{P(E)}, \text{ when } P(E), P(H) > 0.$$

The equation's left side tracks the increase in *H*'s probability brought about by conditioning on *E*. This is one way (among many) to measure the incremental confirmation that *E* provides for *H*. The equation's right side is a way of measuring the marginal change in *E*'s predictability afforded by the supposition of *H*. The theorem thus formalizes the intuition that hypotheses are incrementally confirmed to the extent that their predictions are borne out in experience.

Bayes's theorem reveals many facets of this *evidence–prediction duality*. For example, it relates *odds ratios* to *likelihood ratios*. The *odds* of one hypothesis *H* relative to another *H** is the ratio of their probabilities $O(H, H^*) = P(H)/P(H^*)$. Odds conditional on *E* are defined as

$$O_E(H, H^*) = P(H|E)/P(H^*|E).$$

When *P* encodes the total background evidence, the odds ratio $O_E(H, H^*)/O(H, H^*)$ measures the incremental change that *E* makes to the disparity between the total evidence for *H* and the total evidence for *H**. The likelihood ratio $P(E|H)/P(E|H^*)$ is a way of expressing the relative disparity between *H* and *H** in incremental predictive power with respect to *E*. Bayes's Theorem requires the likelihood and odds ratios to coincide:

$$O_E(H, H^*)/O(H, H^*) = P(E|H)/P(E|H^*).$$

So, the degree to which *E* increases the disparity between the evidence for *H* and for *H** always coincides with the disparity between *H* and *H**'s incremental predictive power *vis-à-vis E*.

Probability theory provides many ways to say that conditioning on *E* increases *H*'s probability. Here are four, where $O(H) = O(H, \sim H)$:

	Probability	*Odds*
Incremental	$P(H\|E) > P(H)$	$O_E(H) > O(H)$
Probative	$P(H\|E) > P(H\|\sim E)$	$O_E(H) > O_{\sim E}(H)$

The columns correspond to two (intertranslatable) ways of quantifying uncertainty. The rows represent two ways of thinking about confirmation. *Incremental* relations, which compare unconditional and conditional quantities, concern the degree to which acquiring datum E will perturb the balance of total evidence for H above or below its current value. *Probative* relations compare the *posterior* evidence for H when E is added to the *posterior* evidence for H when $\sim E$ is added. Here the issue is the extent to which the total evidence for H varies with changes in E's probability. When $P(H|E)$ and $P(H|\sim E)$ are close together, changes in $P(E)$ have little effect on $P(H)$, but when they are far apart such changes have a significant impact.

Depending on whether we express disparities in probabilities using ratios or differences, each of these relations gives rise to two confirmation measures:

	Probability	*Odds*
Incremental	$P(H|E)/P(H)$	$O(H|E)/O(H)$
	$P_E(H) - P(H)$	$O(H|E) - O(H)$
Probative	$P(H|E)/P(H|\sim E)$	$O(H|E)/O(H|\sim E)$
	$P_E(H) - P_{\sim E}(H)$	$O(H|E) - O(H|\sim E)$

This is but a small sampling of the measures of evidential relevance that can be defined. They have different formal properties and can seem to deliver incompatible verdicts on particular cases. Consider, for example, the following constraints on confirmation:

> *Law of likelihood*: E supports H more strongly than E supports H^* iff $P(E|H) > P(E|H^*)$.
> *Law of conditional probability*: E supports H more strongly than E^* supports H iff $P(H|E) > P(H|E^*)$.

The first says that the comparative evidentiary import of a single datum for distinct hypotheses is exclusively a matter of the degree to which the datum is predictable on the basis of the hypotheses. The second says that the relative evidential impact of two items of data for a single hypothesis is entirely a matter of the final probabilities of the hypothesis given the data. Some measures satisfy the law of likelihood (e.g., both probability ratio measures), but others violate it (e.g., both odds ratios). Some measures obey the law of conditional probability (e.g., both incremental ratio measures), but others do not (e.g., both probative ratios).

In addition to satisfying different formal properties, measures can seem to disagree about cases. Suppose that Ellen is a randomly chosen citizen of a town inhabited by 990 Baptists, 2 Catholics, and 8 Buddhists. Let H say that Ellen is not a Buddhist. According to all incremental measures, the datum E that Ellen is a Baptist provides exactly the same amount of evidence for H as does the datum E^* that she is a Catholic. The probative measures disagree, saying Ellen's being a Baptist provides a great deal of evidence for H whereas the datum that she is Catholic provides hardly any.

Probabilists draw different morals at this point. Some, e.g. Eells and Fitelson (2002), see the plethora of measures as posing a dilemma. Since the measures are not equivalent, it seems that an adequate quantitative confirmation theory must either choose among them or restrict its scope to cases where all reasonable confirmation measures agree. One might then seek to identify some apparently necessary formal conditions that adequate measures of confirmation must satisfy, and go on to prove that one particular measure satisfies them. Milne (1996) argues for $P(H|E)/P(H)$ in this fashion. Likewise, Eells and Fitelson (2000, 2002) appeal to formal considerations, including the law of conditional probability, to rule out measures other than (log of) the incremental odds ratio. Alternatively, one might despair of finding any one correct measure and adhere only to claims about confirmation that are invariant under all reasonable measures.

A third approach, advocated by Joyce (1999, 2004), denies that there is any problem. Rather than being competitors, the various measures capture distinct, complementary notions of evidential support. Recall Ellen. When the incremental measures say that E and E^* provide equal evidence for H, this means only that both items of data increase the total evidence for H by the same increment, $1 - P(H)$. When the probative measures say that E is better evidence than E^* is for H, this means that the total evidence for H, as it currently stands, depends much more on information about E's truth-value than on information about E^*'s truth-value. (The disparity between $P(H|E) = 1$ and $P(H|{\sim}E) = 0.2$ far exceeds the disparity between $P(H|E^*) = 1$ and $P(H|{\sim}E^*) = 0.99198$.) When understood this way, these claims clearly do not conflict.

The distinction between incremental and probative evidence dissolves other issues in probabilistic confirmation theory. Take the *problem of old evidence* (Glymour 1980): explaining how someone who is certain or nearly certain of E, and who knows that H entails E, can see E as evidence for H. Highly probable evidence often seems to have great evidentiary value even when the values of $P(E)$, $P(E|H)$ and $P(E|{\sim}H)$ are nearly identical, thus preventing any of the incremental measures of evidence from being large. For example, when Einstein recognized that his new hypothesis of General Relativity entailed the well-known anomalous advance of Mercury's perihelion, he saw this "old evidence" as strongly supporting his theory. As Christensen (1999) and Joyce (1999) suggest, the problem evaporates once we countenance more than one probabilistic notion of evidential support. Antecedently probable data cannot have much *incremental* effect since they are already incorporated into the total evidence. They can, however, still have great probative value: the total evidence for a hypothesis can vary greatly depending on the data's probability.

The contrast between incremental and probative evidence can be made more vivid by the following principle:

> *Surprisingness*: For fixed values of $P(E|H)$ and $P(H)$ with $P(E|H) > P(E)$, the degree to which E confirms H decreases with increases in $P(E)$.

This is a precise formulation of the oft-heard idea that, *ceteris paribus*, hypotheses are better confirmed by unlikely data than by likely data. Surprisingness is not, however, an incontestable fact about confirmation: many philosophers have held that the prior probability of data is irrelevant to their confirming power – see Hempel (1966: 38). And people with disparate opinions about the probability of data often agree about central aspects of their evidential significance. For example, on the basis of preliminary examinations, one clinician might be almost certain that Josh has strep throat, while another might deny this. The clinicians will also disagree about the probability of a strep test on Josh yielding a positive result. But, even though the incremental effect of the test data will be different for each clinician (in virtue of its different surprisingness for them), they can still agree about the data's probative value: both recognize that a positive result, expected or not, will leave the hypothesis well supported, while a negative result will leave it poorly supported.

How is P to be interpreted?

Any assessment of probabilistic confirmation theory must depend on the nature of the probability functions that underlie the enterprise. Various interpretations might be given to *P*. On a subjectivist "Bayesian" reading, *P* captures the strengths of somebody's opinions: probabilistic confirmation theory concerns the doxastic states of individuals. Many object to the use of subjective probabilities in confirmation theory on the grounds that an individual's credences have no place in science, since they are a function both of her prior personal judgments and biases and the particular sequence of evidence she happens to receive (see, e.g., Sober 2002).

In response, Bayesians often observe that the subjectivity of a probability does not render it inaccurate or ill-founded. Credences of competent scientists are excellent guides to the truth in most areas of inquiry. Bayesians sometimes seek to buttress these remarks with "convergence theorems" which show that, under certain conditions, idiosyncratic differences in priors will tend to "wash out" as the evidence increases, thus making the probabilities more "objective"

But some probabilists want more, and aim to provide *P* with an objective interpretation that does not depend on what anyone happens to believe. The most influential attempt to do this, in philosophical circles, is Carnap's.

Logical probability: Carnap's program

The *logical interpretation* of probability seeks to determine universally the degree of confirmation that evidence *E* confers on hypothesis *H*. Pioneered by Johnson and Keynes, and developed most fully by Carnap, the goal is to provide an *inductive logic* that generalizes entailment to *partial entailment*.

Carnap's early (1950) systems begin with a first-order language containing a finite number of monadic predicates and countably many individual constants. The most detailed descriptions in the language – *state descriptions* – affirm or deny the attribution of each predicate to each individual. For example, in a language containing the predicate "*F*" and the constants "*a*," "*b*," and "*c*," the state descriptions are:

1 $Fa \,\&\, Fb \,\&\, Fc$ 2 $Fa \,\&\, Fb \,\&\, \neg Fc$
3 $Fa \,\&\, \neg Fb \,\&\, Fc$ 4 $\neg Fa \,\&\, Fb \,\&\, Fc$

5 $Fa \,\&\, \neg Fb \,\&\, \neg Fc$ 6 $\neg Fa \,\&\, Fb \,\&\, \neg Fc$
7 $\neg Fa \,\&\, \neg Fb \,\&\, Fc$ 8 $\neg Fa \,\&\, \neg Fb \,\&\, \neg Fc$

The choice of a probability measure m for state descriptions induces a *confirmation function*:

$$c(H, E) = \frac{m(H \,\&\, E)}{m(E)} \; (m(E) > 0).$$

A *structure description* is a disjunction of state descriptions that agree on how many individuals instantiate each predicate. For example, the disjunction of state descriptions 2, 3, and 4 yields the structure description characterized as "two F's, one ¬F." Carnap's preferred measure, m^*, gives equal weight to each structure description, these weights in turn shared equally among the constituent state descriptions. In our example, there are four structure descriptions, corresponding to

"three F's,"
"two F's, one ¬F,"
"one F, two ¬F's,"
"three ¬F's."

They each receive 1/4 of the probability, subdividing it equally internally. Thus, m^* assigns 1/4 to state descriptions 1 and 8, and 1/12 to the rest. In contrast to c, the resulting confirmation function c^* allows inductive learning: evidence of some individuals' having a property confirms other individuals' having that property. For instance, the *a priori* probability of Fa is $m^*(Fa) = 1/2$. However,

$$c^*(Fa, Fb) = \frac{c^*(Fa \,\&\, Fb)}{c^*(Fb)}$$
$$= \frac{1/3}{1/2}$$
$$= 2/3.$$

So, the evidence that Fb confirms the hypothesis that Fa.

While the early Carnap favored c^* for its simplicity and salience, it is not obvious that it is the unique confirmation function he sought, since infinitely many candidates have this "inductive learning" property. He later (1962) generalizes his confirmation function to a continuum of functions c_λ. He considers languages containing sets of one-place predicates such that, for each individual, exactly one member of each set applies. He lays down a host of axioms of symmetry and inductive learning. They imply that, for the set of predicates $\{P_i\}$, $i = 1,2,\ldots,k$, $k > 2$,

c_λ (individual $n+1$ is P_j, n_j of the first n individuals are P_j)

$$= \frac{n}{n+\lambda}\left(\frac{n_j}{n}\right) + \frac{\lambda}{n+\lambda}\left(\frac{1}{k}\right), \text{ where } 0 < \lambda < \infty.$$

The bracketed fractions are respectively the proportion of observed "successes," and the symmetrically assigned *a priori* probability; their unbracketed weights sum to 1. λ is an index of "caution": the higher it is, the less responsive is c_λ to evidence. At $\lambda = 0$, we have the inductively incautious "straight rule" that simply equates the conditional probabilities to the corresponding relative frequencies. At $\lambda = \infty$, we have the rigid method that never learns from experience. In between we have the range of all admissible inductive rules. Carnap regards the choice of λ as a pragmatic matter, something to be decided in a particular context.

Several problems for Carnap concern the languages over which his confirmation functions are defined. These languages are clearly too impoverished to do justice to much scientific theorizing; yet as they are enriched with further expressive power, the confirmation relations change. Still more seriously, these relations are determined solely by the *syntax* of the sentences – their *meanings* play no role.

The fact that meanings should play a role is one lesson of Goodman's *new riddle of induction*. Our evidence of observing many green emeralds surely confirms that emeralds observed at any future time will be green. Now consider the predicate "grue," which applies to objects that are green and observed before some future time *t*, or blue and not observed before *t*. Our evidence can be equivalently described as the observation of many *grue* emeralds; but it does *not* confirm that emeralds observed after *t* will be grue – for that would mean that they are blue. The challenge for any confirmation theory is to account for the differing confirmation relations that our evidence bears to the "green" and the "grue" hypotheses. Any such theory must apparently be sensitive to features beside syntactical form, since syntactically "green" and "grue" are on a par.

One might protest that "grue" is somehow syntactically more complex than "green" – after all, "grue"'s definition above involves a somewhat complicated disjunction. But now define "bleen," which applies to objects that are observed before *t* and blue, or not observed before *t* and green. Then there is an alarming interdefinability of the "green/blue" and the "grue/bleen" vocabulary. In particular, an emerald is green iff it is grue and observed before *t*, or bleen and not observed before *t*. So what counts as a "complicated disjunction" depends on which predicates we start with. Nor will it help to claim that "grue" is in some sense "gerrymandered," or "positional" (referring as it does to a particular time, *t*). For whatever these pejoratives may mean, the interdefinability point will underwrite the same claims about "green."

So Carnap's languages apparently have to privilege certain predicates over others – presumably outlawing monstrosities such as "grue." It is hard to see how this can have any basis in *logic*, and how this privileging can be done in a principled way. Goodman, for example, appeals to the somewhat nebulous notion of *entrenchment:* a predicate is entrenched iff we have used it in successful inductive inferences in the past. But our commonsense predicates are often better entrenched than those of science; that is hardly a reason to favor the former when making predictions.

Finally, return to the dependence of Carnap's confirmation functions on the parameter λ. Nothing in logic determines, or even constrains, its value. Carnap thought that it might be determined empirically, but the bearing of empirical data on its value is itself a problem of confirmation, and an infinite regress threatens. This problem is only exacerbated for the late Carnap (1971), when he generalizes his system further to include analogical considerations. This involves a further parameter over whose setting there is again much freedom, and certainly no constraint from logic. We have thus come a long way from his initial hope for a unique confirmation function. A slightly different, and increasingly popular, "objectivist" interpretation of probability is found in the MaxEnt or "objective Bayesian" program of E. T. Jaynes (2003). For details, see Colin Howson's article "Bayesianism" in this volume.

Conclusions

We began by noting how little of our reasoning is captured by deductive logic, and how there is an apparent need for confirmation theory. Carnap's *inductive logic* was intended to assimilate confirmation theory to deductive logic. To be sure, confirmation theory does bear some interesting analogies to deductive logic: it is not a matter of the truth of some piece of evidence *E*, nor of some hypothesis *H*, but rather of the bearing that *E* has on *H*. But we have learned that there are apparently some important disanalogies. Unlike deductive entailment,

- Confirmation relations come in varying degrees.
- The relations cannot be captured purely syntactically: meanings of terms are important.
- The relations may not be uniquely constrained.
- They apparently involve at least a three-place relation, between an evidence sentence *E*, a hypothesis *H*, and background knowledge *K* (which may be captured in a probability function *P*).

That said, we side with Carnap, and against Hume and Popper, in insisting that relations of confirmation may be non-trivial, of importance to science, philosophy, and daily life, and susceptible to genuine illumination.

Acknowledgements

We thank Martin Curd, Franz Huber, and Stathis Psillos for very helpful comments on earlier drafts.

See also Bayesianism; Evidence; Prediction; Probability; Scientific method; Underdetermination.

References

Ayer, A. J. (1936) *Language, Truth and Logic*, London: Penguin.
Carnap, R. (1962) *Logical Foundations of Probability*, 2nd edn, Chicago: University of Chicago Press.

Christensen, D. (1999) "Measuring Confirmation," *Journal of Philosophy* 96: 437–61.
Duhem, Pierre (1905) *La Théorie physique: son objet, sa structure*, 2nd edn, Paris: Marcel Rivière, 1914; trans. P. P. Wiener as *The Aim and Structure of Physical Theory*, Princeton, NJ: Princeton University Press, 1954.
Earman, J. (1992) *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, MA: MIT Press.
Eells, Ellery and Fitelson, Branden (2000) "Measuring Confirmation and Evidence," *Journal of Philosophy* 97: 663–72.
—— (2002) "Symmetries and Asymmetries in Evidential Support," *Philosophical Studies* 107: 129–42.
Fitelson, Branden (2002) "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction," *Philosophy of Science* 69: 611–22.
Glymour, Clark (1980) *Theory and Evidence*, Princeton, NJ: Princeton University Press.
Good, I. J. (1967) "The White Shoe Is a Red Herring," *British Journal for the Philosophy of Science* 17: 322.
Hempel, Carl (1945) "Studies in the Logic of Confirmation," *Mind* 54: 1–26, 97–121.
—— (1966) *Philosophy of Natural Science*, New York: Prentice-Hall.
Jaynes, E. T. (2003) *Probability Theory: The Logic of Science* (ed. by G. Larry Bretthorst), Cambridge: Cambridge University Press.
Johnson, W. E. (1932) "Probability: The Deductive and Inductive Problems," *Mind* 41: 409–23.
Joyce, James M. (1999) *The Foundations of Causal Decision Theory*, Cambridge: Cambridge University Press.
—— (2004) "Bayesianism," in A. Mele and P. Rawling (eds) *The Oxford Handbook of Rationality*, Oxford: Oxford University Press, pp. 132–55.
Keynes, John Maynard (1921) *A Treatise on Probability*, London: Macmillan.
Milne, Peter (1996) "Log[$p(h/eb)/p(h/b)$] is the One True Measure of Confirmation," *British Journal for the Philosophy of Science* 20: 21–6.
Moretti, Luca (2003) "Why the Converse Consequence Condition Cannot Be Accepted," *Analysis* 63: 297–300.
Quine, W. V. (1951) "Two Dogmas of Empiricism," *Philosophical Review* 60: 20–43.
—— (1969) "Natural Kinds," in W. V. Quine, *Ontological Relativity and Other Essays*, New York: Columbia University Press, pp. 114–38.

Further reading

Two excellent overviews of Bayesian confirmation theory are: Colin Howson and Peter Urbach, *Scientific Reasoning: The Bayesian Approach*, 3rd edn (La Salle, IL: Open Court, 2005); and John Earman, *Bayes or Bust?* (Cambridge, MA: MIT Press, 1992). Clark Glymour (1980) advocates his "bootstrap" theory of confirmation. John Earman (ed.) *Testing Scientific Theories* (Minneapolis: University of Minnesota Press, 1983) is a volume of responses. Useful discussions of the problem of old evidence can be found in: Lyle Zynda, "Old Evidence and New Theories," *Philosophical Studies* 77 (1995): 67–95; David Christensen, "Measuring Confirmation," *Journal of Philosophy* 96 (1999): 437–61; and Jim Joyce, *The Foundations of Causal Decision Theory* (Cambridge: Cambridge University Press, 1999). The problem of measuring confirmation is discussed in many of Branden Fitelson's papers, available at http://fitelson.org/research.htm. Recent defenses of the law of likelihood can be found in Richard M. Royall, *Statistical Evidence: A Likelihood Paradigm* (New York: Chapman & Hall, 1997) and papers by Elliott Sober available at http://philosophy.wisc.edu/sober. A detailed discussion of the interpretations of probability, including the logical interpretation, can be found in Alan Hájek, "Interpretations of Probability," in Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (summer 2003 edition); available online: http://plato.stanford.edu/archives/sum2003/entries/probability-interpret. Jon Williamson's *In Defence of Objective Bayesianism* (Oxford: Oxford University Press, 2010) provides a spirited recent defense of objective Bayesianism. Psychologists have recently become interested in human tendencies to judge certain experiences confirmatory and disconfirmatory; see for example Tentori, Katya, Vincenzo Crupi, Nicolao Bonini, and Daniel Osherson, "Comparison of Confirmation Measures," *Cognition* 103 (2007): 107–19 and Tentori, Katya, Vincenzo Crupi, and Daniel Osherson, "Determinants of Confirmation," *Psychonomic Bulletin and Review* 14 (2007): 877–83.