Approximations to Truth: Confirmation and Verisimilitude (4)

If inquiry aims at the truth, what should inquirers be doing? Two aspects to this: attitude, content. If A is true, our attitude towards it should be as close as possible to full belief. If A is our theory, it should be as truthlike as possible. Confirmation theory is directed at the first goal. The theory of verisimilitude is directed at the other.

QUALITATIVE CONFIRMATION Hempel sees confirmation theory developing in three stages: qualitative, comparative, and quantitative. Qualitative is supposed to satisfy

Entailment E confirms any H that it entails.
Consistency If E confirms H, it does not confirm anything contradicting H.
Special Consequence If E confirms H, it confirms what H implies.
Converse Consequence If E confirms H, it confirms whatever implies H.

Against the last, he notes that Rudy is black would confirm Rudy is black & F = ma. Also that adding it to the others trivializes the confirmation relation. P confirms P, so it confirms P&Q by converse consequence, so it confirms Q by special consequence.

QUANTITATIVE CONFIRMATION Epistemologists have increasingly wanted to turn Hempel’s three stages around...to define qualitative and comparative confirmation in probabilistic terms. The turning point was Carnap’s critique of Hempel’s conditions.

Hempel is running two kinds of confirmation together, Carnap thinks, whose differences emerge only when we consider the matter quantitatively. E confirms H incrementally if p(H|E) > p(H). It confirms H absolutely if p(H|E) is close enough to 1.

These should not be confused!! But is Hempel confusing them? You’d expect Carnap to argue that some of them hold only for the absolute notion, others only for the incremental. But the truth is that all of Hempel’s conditions hold for the absolute notion! It is only converse consequence that fails to hold absolutely—and he rejects it.

But, Hempel’s rhetoric and examples suggest the relative or incremental notion. A black crow makes it likelier, not absolutely likely, that all crows are black. Incremental confirmation meets neither of Hempel’s key conditions. Not consistency, because Rudy is black relatively confirms both of Rudy is black and poetic, Rudy is black and prosaic. Not special consequence, since Rudy is black boosts the chances of Rudy is a black and poetic but not Rudy is poetic. Carnap is right. Hempel’s conditions do not hold of the notion of confirmation he had in mind—or at least of the one he had more in mind of the two Carnap offers him. This leaves an opening, to which we return.

there may be some third probabilistic [notion of] confirmation that allows Hempel...to pass between the horns of this dilemma. But it is up to the defender of Hempel’s...[approach to] confirmation to produce the tertium quid” (Earman [1992a]).

SURPLUS CONTENT Suppose H implies E. Then one kind of confirmation obtains automatically. E rules out certain ways H might have been wrong: the ones implying ¬E. I was born on a Thursday supports Everyone was born on a Thursday by eliminating one possible counterexample.

A more interesting way is for an entailment E of H to bear favorably on H is to support the rest of H—its surplus content relative to E. I was born on a Thursday does not bear favorably in this way on Everyone here was born on a Thursday; it lends no support to the untested part. The distinction is between two kinds of incremental confirmation: “mere content cutting” (Gemes [1994]) and, to give it a name, inductive confirmation.

Inductive confirmation is defined in terms of surplus content; views about what H adds to E will guide one’s thinking about when inductive confirmation occurs. Popper and
Miller claims in a 1983 letter to *Nature* that it never occurs. Yes, $H$ is logically conditional on $E$ than on its own; $E$ confirms $H$ in the sense favored by Bayesians. The added probability does not make for *inductive* confirmation, though, as we see on isolating $H$’s surplus content.

1. $H$ is logically equivalent to $H \lor E$ & $H \lor \neg E$.
2. $E$ entails $H \lor E$; so the surplus content in that conjunction is $H \lor \neg E$.
3. $E$ makes $E \supset H$ less likely: $p(H \lor \neg E \mid E) < p(H \lor E)$, provided $p(H \mid E) \neq 1 \neq p(E)$.
4. $E$ does not inductively confirm $H$.

The step from 1 to 2 is based on what theory of surplus content? Apparently this: if $H$ is equivalent to the conjunction of $C$ with (independent?) $D$, and $E$ implies $C$, then the surplus content is $D$. But that underdetermines surplus content badly. Suppose $H$ is $P \land Q$. Popper will factorize it as $(P \lor P \land Q) \land (\neg P \lor P \land Q)$, which suggests $H$’s surplus content over $P$ should be $\neg P \lor P \land Q$. But one could just as easily treat it as the conjunction it is, in which case $H$’s surplus content over $P$ is $Q$. Now the step from 2 to 3 doesn’t go through; it depended on Popper’s factorization which is just one among many.

Is $H \lor \neg E$ somehow first among equals? Not that either. $H \lor \neg E$ is $E \supset H$, whose candidacy for the role of $H-E$ has already been rejected. As Richard Jeffrey says in his response to Popper: the surplus content of *Everyone’s thirsty* over *Joe is thirsty* is not *Everyone is thirsty if Joe is*, but *Everyone is thirsty with the possible exception of Joe*.

A THIRD WAY Hempel needs “a third probabilistic notion of confirmation” to respond to Carnap’s critique. Inductivists need one to respond to Popper’s.

1. $E$ fully confirms $H$ if it bears favorably on $H$ and all its parts.

Bearing-favorably-on can be just what Carnap and Popper say: probabilification. Full confirmation is a turbocharged development of this: *pervasive* probabilification.

2. $E$ pervasively probabilifies $H$ iff for all (nontrivial) $F \leq H$, $p(F \mid E \& K) > p(F \mid K)$.

Full confirmation and *Consistency*: Rudy is black & happy and Rudy is black & unhappy are incrementally confirmed by Rudy is black, but not fully confirmed—pervasively probabilified—by it. To fully confirm them, it would have to make it likelier both that Rudy is happy and that he isn’t. *E* cannot pervasively probabilify $H$ and $H'$ if they have $F$ and $\neg F$ (resp.) as parts.

Full confirmation and *Entailment*: Suppose $E$ entails $H$. $E$ entails $F$ (a part of $H$) by transitivity of entailment; so $p(F \mid E) = 1$. $E$ Unless $p(F)$ is 1 already, $E$ probabilizes it.

Call $H$ *new* if $p(F)$ was not 1 already, for any nontrivial part $F$ of $H$. Evidence entailing a new hypothesis pervasively probabilifies that hypothesis.

Full confirmation and *Special consequence*: An $E$ that fully confirms $H$ may not fully confirm $H$’s consequences. Rudy the black crow fully confirms *Crows are black*, but not *Crows are black, if Rudy is black*. But, is $E \supset H$ really the kind of thing Hempel means to share in the confirmation of $H$? He introduces the consequence condition in this passage:

> [A consequence $G$ of $H$] is but an assertion of *all or part* of the ... content of $H$ and has therefore to be regarded as confirmed by any evidence which confirms $H$ (SLC II, 103).

He seems to be saying that $E$’s confirmation of $H$ should penetrate down to $H$’s parts. So it does on our model: if $E$ p-probabilifies $H$, then by transitivity of containment, it p-probabilities any $F$ that’s contained in $H$.

Hempel’s conditions do hold, more or less, for full or “pervasive” confirmation, which is incremental as he intends. Is it rewriting history, to attribute to him an interest in full
confirmation? Not attributing it might be rewriting history! Recall his objection to converse consequence: *Rudy is black* does not confirm that *Rudy is black & happy*. Rudy’s blackness does confirm that conjunction in the sense of probabilifying it. What it doesn’t do is probabilify the conjunction’s parts. Hempel’s objection rests on an obviously false assumption, if he is not thinking of confirmation as full or pervasive.

Hempel’s own model is this. *H* is confirmed by its “development” *H|I* for observed individuals *I*. This ought intuitively to be what *H* says about observed individuals. It’s defined, though, as what *H* would say if nothing existed but the members of *I*. This is sometimes stronger than *H*; *∃xFx|I* says that observed *Fs* exist. And it’s often irrelevant to *H*. Monism is not confirmed by the fact that it is true of observed individuals, when only one thing has been observed. Hempel may have mean to say this: *H* confirmed by the part of *H* that concerns some pertinent subject matter, say, actual observations or things already known. His theory so understood has something common with the hypothetico-deductive model, which is discussed next.

**HYPOTHETICO-DEDUCTIVISM** Quantitative confirmation has won the day. Bayesians do like, however, to cast a glance back at the qualitative literature, if only to take a victory lap around it before getting down to business. Not everything in that literature can be saved. But this is as it should be. What was right in the qualitative tradition is explained by Bayes, we think, and what was wrong is refuted by Bayes.

Take the HD model of confirmation. One tests a hypothesis, on that model, by seeing whether its predictions—really, consequences—check out. False consequences count against *H*; true ones count in its favor; they confirm it. Suppressing all complexities,

\[(HD) \ E \text{ confirms } H \text{ if } H \text{ implies } E.\]

This is expected, if confirmation is favorable probabilistic relevance. From *H* entailing *E*, it follows that \(p(H|E) \geq p(H)\). This is one of a traveling cast of Bayesian “success stories” trotted out in the opening chapters of textbooks. The form is: probability calculations let us derive principles that others have had to treat as basic. But, some of the vindicated principles have been questioned.

**Tacking by Conjunction** If *H* entails *E*, then *H&E* entails *E* too. The DN model predicts, and the Bayesian backs it up on this, that the class of items confirmed by a piece of evidence is closed under the operation of tacking on random conjuncts. *Rudy is black* confirms *All crows are black*, yes, but also *All crows are black & everything else is white*. *Rudy is black* & *Rudy is happy* is confirmed not only by *Rudy is black*, but also *Rudy is black ∨ All crows are not black*.

What now? There is something right in the idea that a theory is confirmed by its consequences. But it can’t be maintained in full generality. For a theory to be confirmed by any old consequence runs into the tacking by disjunction problem. For its consequences to confirm any old theory runs into the tacking by conjunction problem. If *E* is the right kind of *H*-consequence, one nevertheless feels, and *H* is the right kind of *E*-implier, the relation should hold (Gemes [1998]).

The wrong kind of *H*-consequence is the kind that tacks on a random disjunct. That’s the paradigm of a “mere” consequence. The obvious solution is to say that *H* is HD-confirmed only by those of its consequences that are parts.

The wrong kind of *E*-implier is the kind that tacks on a random conjunct. *All crows are black & The Earth moves* is not confirmed by *Rudy is black*. The obvious solution is to say that all of *H*’s parts are called on in the derivation of *E*. But, *E* is itself a part of *H* (we just said) so the other parts need not be called on. Maybe all the theory’s axioms are

“[O]ne may well be inclined to agree that a generalization such as ‘All emeralds are grue’ is not lawlike, and that its applicability to as yet unexamined cases is not attested to by its previously established instances” (Hempel [1960]). A lawlike generalization then presumably should have the full confirmation property.

Howson and Urbach [1989], Earman [1992b].
called on? That asks too much, since All crows are black can be rendered as \(\{\text{All crows are black, Rudy is black}\}\). Maybe the axioms have to be non-overlapping? Well, but the members of \(\{\text{All crows are black if Rudy is, Rudy is black}\}\) are non-overlapping.

Proposal, due to Gemes: \(T\) is naturally axiomatized by \(A_1 \ldots A_n\), iff (i) no \(A_i\) is a “mere consequence” of \(T\)—each is a part, and (ii) no part of any \(A_i\) is entailed by the others (Gemes [1993]). A theory is (fully) DN-confirmed by those of its parts whose derivation call on each axiom, for every natural axiomatization.

THE RAVENS This talk of full or pervasive confirmation puts us in mind of enumerative induction and Hempel’s paradox of the ravens. It has four elements: three plausible-looking premises and a nutty-looking apparent consequence of those premises.

Nicod’s Criterion (NC): All Fs are Gs is confirmed by its instances

Equivalence Condition (EC): Evidence for \(H\) is evidence for its logical equivalents.

Equivalence Fact (EF): All Fs are Gs is logically equivalent to All non-Gs are non-Fs

Paradoxical Conclusion (PC): Ravens are black is confirmed by non-black non-ravens.

The standard Bayesian story just embraces PC. A non-black non-raven does confirm—incrementally—that all ravens are black. But, it confirms it just the teeniest little bit—not as much as a black raven does. The idea was apparently first suggested in 1940 by the Polish logician Janina Hosiasson-Lindenbaum (Hosiasson-Lindenbaum [1940]). A randomly chosen item is likelier non-black than a raven, hence we sample a larger portion of the counterexample space by looking at ravens.

The problem was noted already by Hempel. "[I]s this last numerical assumption [that non-black things greatly outnumber ravens] warranted in the present case and analogously in all other "paradoxical" cases?" Hempel seems to be suggesting that the paradox still arises if a randomly chosen item is just as likely a raven as non-black.

This is hard to imagine, so consider a different case. \(H\) is a lawful generalization to the effect that Charged particles have no spin—they’re “rotationally inert,” for short, inert.

The numerical assumption becomes: a randomly chosen particle is likelier spinnier than charged. This might well be false; there might be as many charged particles as spinnier ones. Doesn’t it still seem that a charged inert particle confirms Charged particles are inert more, or better, than a spinnier neutral one?

The initial generalization is about charged particles; how could a neutral particle tell us about them? The most it can hope to do is block a potential counterexample. Spinny particles are neutral might be a lawful generalization in its own right. The most a particle’s spinlessness can do is block a potential counterexample: it at any rate is not spinnier and charged. They differ confirmationally, because mereologically; they differ mereologically because they differ in what they’re about.

TRUTH & CONFIRMATION Wait! Confirmation is to do with likelihood of truth. Fs are G and Non-Gs aren’t F are true in the same circumstances, hence equiprobable. How could \(E\) confirm one more than the other? Two worries can be distinguished here.

Philosophical worry. Why call it confirmation if it doesn’t line up with probability of truth? Well, but it does. All Fs are Gs is not the only truth in the neighborhood: Fs are thereby Gs, Fs would still be Gs if more things were F, Fs are Gs as a matter of law. For short, \(\Box_F\) (Fs are Gs). Yes, Fs are Gs is true in the same worlds as non-Gs are non-Fs. \(\Box_{\sim G}\) (Fs are Gs) and \(\Box_{\sim F}\) (non-Gs are non-Fs) are another matter.

Technical worry. Rudy supposedly confers likelihood on the parts of Ravens are black, but not the parts of Non-black things are non-ravens. How can this be when they’re constrained to be equiprobable? Whatever drives the one’s probability up, drives the other’s up equally. This creates pressure for \(H\) and its contrapositive to have the same or cor-

What about \(\{\text{All observed crows are black, All unobserved crows are black}\}\)? That raises deeper issues. The HD model is not going to be much use against fractured, grue-like, decompositions.

Hempel [1945a,b]
responder parts, contrary to the suggestion above that the two generalizations differ confirmationally because mereologically.

The answer will be that the corresponding parts, although equal in probability, differ in their relation to the whole. “Real” parts of All ravens are black correspond to degenerate parts of its contrapositive. Rudy figures in what is said by the first, while it limits what is said by the second.

What is the logical form of All ravens are black? Is it (as Hempel thought) ∃x(Rx ⊃ Bx), that is, ∃x (Rx&¬Bx)? That would make my non-raven-hood a factor in H’s truth. Intuitively it’s a factor rather in what its truth requires.

Conditional assertion approach (Belnap [1970]). Instead of Rx ⊃ Bx, put Rx//Bx, where to assert Rx//Bx is to assert (but only if x is a raven) that x is black. This is logically revisionary, though; A ⊃ C is equivalent to ¬C ⊃ ¬A, but the same cannot be said of asserting C, conditionally on A, and asserting ¬A, conditionally on ¬C.

Better: Bx provided that Rx. Rx:Bx is true in the same worlds as Rx ⊃ Bx. But the reasons differ. One is true because x is either not a raven or black. The other is true, should x be a raven, because x is black. Otherwise it is trivially true—true without benefit of truthmaker, ∀x(Rx ⊃ Bx) is true because everything is either not a raven or black. ∀x(Rx: Bx) is true because certain things are black, viz. the ravens.

The probability of Fa:Ga can be raised in two ways. One kind of evidence lowers p(Fa), thus boosting the chances of Fa:Ga being trivially true. The other lowers p(Fa&¬Ga) (leaving p(Fa) unchanged), boosting the chances of Fa:Ga being substantively true. Similarly at the level of generalizations. One kind of evidence increases p(∀x(Fx:Gx)) by making more of its content trivial—by cutting into what it (non-trivially) says. The other does it by making the non-trivial part(s) more probable.

To the technical worry, we reply that each Rx:Bx has a counterpart ¬Bx:¬Rx that agrees with it in probability, but with substantive and trivial probability interchanged. As the chances rise of Rudy being (substantively) black if a raven, they rise of Rudy being (trivially) a non-raven if non-black. The process repeats at the level of generalizations: Rudy’s effects on the probabilities of ∀x(Rx:Bx) and its contrapositive are the same. The difference is only that it confirms what the first says, while reducing what it says.

VERISIMILITUDE Beliefs should be close to the truth; confirmation theory is supposed to help us achieve this. But closeness to the truth has an aspect that confirmation theory is blind to....we want to maximize the amount of truth we believe and minimize the amount of falsehood. Popper was famously pessimistic about the first aim; he tended to emphasize the second. Science progresses, not when our theories are better confirmed, but when they achieve greater verisimilitude. His initial definition, with Θ and Σ ranging over theories:

Θ’s truth-content Θ^T is the set of its true consequences.
Θ’s false-content Θ^F is the set of its false consequences.
Σ is more truthlike than Θ iff Θ^T ⊆ Σ^T and Σ^F ⊆ Θ^F....and not conversely.

Some nice consequences, e.g., among true theories, verisimilitude goes with strength. The problem is on the false side. If Θ and Σ are false and independent, each has truth-content lacked by the other; Θ alone implies Σ^T ⊃ Θ^T, Σ alone implies Θ^T ⊃ Σ^T (Tichy-Miller).
False theories are left completely unranked by Popper’s proposal.

Attention turned from the content approach to the likeness approach: Θ has greater verisimilitude to the extent it holds in worlds closer to actuality. In the diagram, C is closest, to the truth, B is furthest, and A and D are incomparable. But there is no simple answer to how the distances of individual Θ-worlds from @ are to be combined into a single measure suitable for the set of them. There may for Arrow’s theorem reasons be no good way of doing it (Zwart and Franssen [2007]).
Back to basics. Popper went wrong in equating truth-content with true consequences. What if Θ’s truth-content = the set of its true parts. Σ is more truthlike than Θ iff its true parts imply Θ’s true parts, and Θ’s false parts imply Σ’s (and not conversely).

Not quite: suppose Σ = P&Q, Θ = Q, where P is true and Q is false. Σ should come out ahead since it adds a true conjunct. But it has a false part not implied by Θ, namely itself. This seems a cheat; what’s false about P&Q, namely Q, is part of Θ. Switch to Σ’s wholly false parts—the ones with no non-trivial true parts buried within. (True parts are wholly true automatically).

Σ is more truthlike than Θ iff

Σ’s wholly true parts imply Θ’s, and its wholly false parts are implied by Θ’s.

This makes for a further form of hyperintensionality; logically equivalent hypotheses can differ in verisimilitude. *I am mortal* has less of it than *All men are mortal*, but not than *Immortals are never human*. *I am mortal, if human* is a truth contained in *All men are mortal* that is not implied by truths contained in *Immortals are never human*.

Logical subtraction has a role to play in confirmation theory, via the notion of surplus content. Subject matter, too; it lets us reply to the paradox of the ravens that logically equivalent generalizations can be about different things, which affects their evidential relations. Content mereology lets us define pervasive probabilification, which helps with the tacking paradoxes. Inductive skeptics like Popper don’t care about confirmation. But they derive some advantage too; we need content-part to explain what it is for one theory to have more truth in it, and less falsity, than another.

References


