Genuine Confirmation and the Use-Novelty Criterion

Gerhard Schurz (Universität Düsseldorf)

1. The Problem: Bayesian Confirmation of Irrational Beliefs

Neo-Creationists have used Bayesian methods to confirm (*refined* versions of) *creationism* (Swinburne 1979, Unwin 2005)

In contrast to genesis creationism that is falsifiable by its empirical consequences, *refined creationisms are empirically uncriticizable, although they have, logically speaking empirical content*

How possible? By **ex-post constructions**: one enriches the creator hypothesis ex post by scientifically established facts as follows:

Hypothesis of refined creationism (H): Our world has a creator who created it with the following properties: (E) … here follows a list of scientifically established facts.

History of rationalized theology is full of pseudo-explanations of that sort …
→ contemporary *intelligent design* movement (Behe 1996, Dembski 1998)

$\Rightarrow$ According to Bayesian confirmation as well as H-D confirmation: E 'confirms' H.

Note: E confirms all other sorts of 'irrational' explanations, too
(devil, Spaghetti monster …)

With Bayes-confirmation I mean
comparative confirmation of H by E – iff $P(H|E) > P(H)$ iff $P(E|H) > P(E)$
(widely accepted among Bayesians; independent of prior probability of H).

Recall Bayes-formula: $P(H|E) = P(E|H) \cdot P(H) / P(E)$

Bayes-confirmation implies (and the same holds for H-D confirmation):

Bayesian pseudo-confirmation: every contingent hypothesis H that logically entails a contingent true evidence E is confirmed by E. (S is contingent iff $0 < P(S) < 1$)

Can be exploited by speculative thinkers at their pleasure

Bayesians (e.g. Howson/Urbach 1996) counter that scientific hypotheses have a *higher prior probabilit*y than religious speculations,

but that is doubly questionable because:

1) prior probabilities are (more or less) subjective, and

2) it seems that refined creationism is not just a little bit less confirmed than evolution theory, but not confirmed at all.

Conclusion: Bayesian confirmation theory is too weak to *demarcate* genuine confirmation from pseudo-confirmation

→ a demarcation criterion via a notion of genuine confirmability is a *desideratum*…

[because other demarcation accounts fail; cf. *Synthese* 178/2, 2011) …]

2. Alternative Confirmation Concepts: Novel Predictions und Use Novelty

Major characteristics of the above pseudo-explanations:
they are entirely **ex-post** ad-hoc constructions – unable to figure as predictions.

Novel prediction criterion (Musgrave 1974, Lakatos 1977, Ladyman/Ross 2007):
Confirming evidence E must be a novel prediction of the hypothesis H

→ "prediction" is understood here not in the temporal but in the epistemic sense:
E was unknown when H was developed (includes retrodictions; Stegmueller 1983)
→ "novel" means here just "new in the epistemic sense" (stronger notion later)

Objections:
(1) The time when an evidence gets known is subjective (person-relative), while confirmation should be an objective (semantic) relation between propositions.
(2) There exist clear cases of confirmation of scientific theories by evidences that were known long before – e.g. the confirmation of general relativity theory by the deviations of the trajectory of Mercury from classical predictions.

Improvement by Worrall (2010) – *criterion of Use Novelty (UN):*
Confirming evidence E must not have been used in the construction of the hypothesis

⇒ Construction proceeds by fitting a variable parameter x of a more general hypothesis/theory Hx to a special value c, thereby obtaining Hc, a *specialization* of Hx.

	Hx abbreviates $\exists x_1 (x_2\ldots)H[x_1,x_2,\ldots]$	Hc abbrev. $H[c_1, c_2,\ldots]$
E.g.:	God created (variable) facts x.	God created known facts E.

Note: c_i is such a function of E_i such that every possible evidence E_i can be 'ex post' explained by Hc_i.

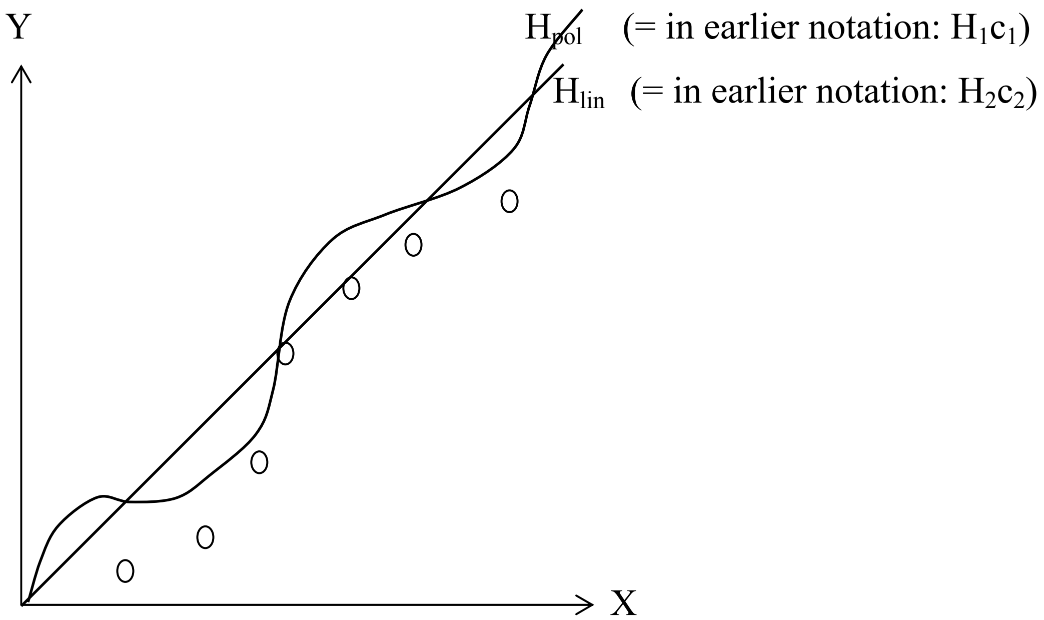
→ I think, Worrall's UN-criterion goes into the right direction. I'll defend it.

3. Curve Fitting – a Paradigm Case for UN

constant ↘ linear ↓ quadratic .... ↙

Polynomial functions: $Y = c_0 + c_1 \cdot X + c_2 \cdot X^2 + \ldots \sigma$ (degree n)

⇒ Every set E of (say) m data points in the X-Y-coordinate system can be approximated by every polynomial function of variable degree up to variable remainder dispersion σ – σ gets smaller the higher the degree of the polynomial and becomes zero $n \geq m+1$.



Linear vs. (high-degree) polynomial curve fitting

H_{pol} approximates the data better than H_{lin}. *Is H_{pol} therefore better confirmed?*

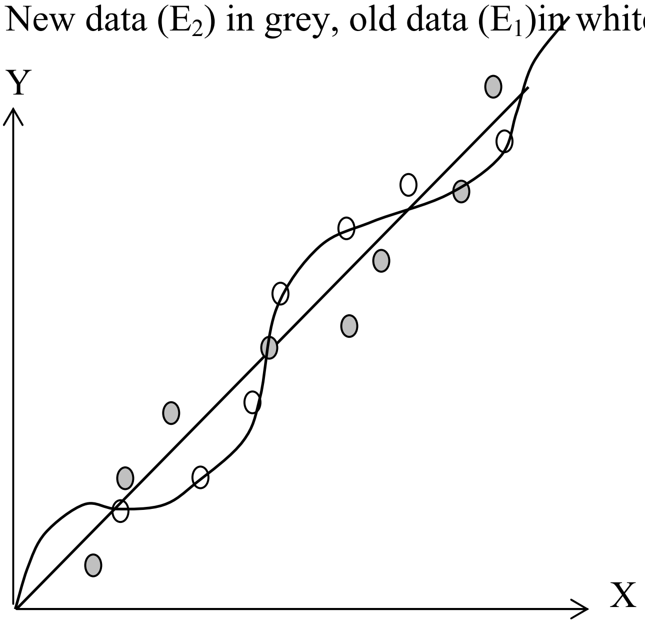
⇒ NO, because of the danger of *overfitting* (fitting on accidentalities of the sample)

Worrall's UN-account fits: both H_{lin} and H_{pol} result from fitting more general hypotheses Lin (=H_1x_1) and Pol (=H_2x_2) to the data set E_1.

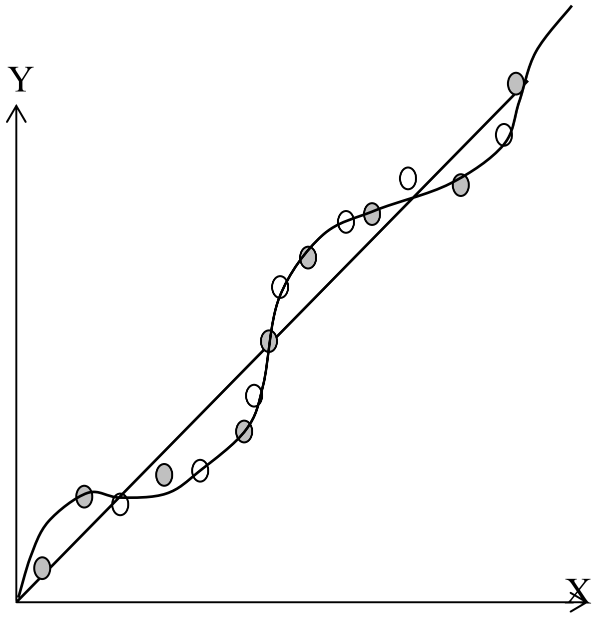
[Formally: Lin = $\exists x_{0,} x_1, \sigma$: $Y = c_0 + c_1 \cdot X + \sigma_1$; likewise for Pol]

Whether H_{lin} or H_{pol} is confirmed can only be seen *at hand of a new data* set E_2 that was not used for fitting the parameters.

New data (E_2) in grey, old data (E_1)in white :



H_{lin} and thus LIN is confirmed by E_2



H_{pol} and thus POL is confirmed by E_2

The UN criterion is a major statistical practice in form of so-called **cross-validation** (Mosier 1951): split a data set D (several times) randomly into D_1, D_2, use D_1 for fitting and test with help of D_2.

Major alternatives: refined criteria that don't apply UN such as AIC and BIC; promoted by Hitchcock/Sober (2004).

⇒ Result of Paulßen (PhD) & Schurz: AIC and BIC are *hopeless inferior* to cross-validation for small data sets:

Number of data points	% mistakes of fitting-method for polynomials 1≤ degree ≤ 15 AIC	BIC	Cross-validation
10	100	100	0
15	100	100	0
20	75	50	0
50	60	10	0
100	25	5	0
500	25	1	0

Explanation: AIC and BIC assign a σ-proportional penalty to more complex hypotheses. This works only if true σ is known, but not when σ is estimated from fitting-result, because overfitting curves underestimate σ.

⇒ *The only safe guard against overfitting are tests with new data sets!*

Objections to Worrall's UN -account:

(1) Worrall's claims that the implication Hx→Hc is logically entailed by E.

⇒ This is sometimes but *not always* true. In the case of curve fitting Hx→Hc is merely *inductively* confirmed by E (by the statistical maximum likelihood criterion)

More difficult objections:

(2) The UN criterion doesn't apply to simple inductive confirmations:
hypothesis about the domain-frequency is confirmed by a sample-frequency, although it was obtained by fitting (Howson 1990, Mayo 1996)

Worrall's reply (2010, p. 69f): not a representative case of confirmational tests (?)

(3) Also the UN-criterion is subjective (person-relative), because different scientists may arrive at the same hypothesis along different routes (Musgrave 1974).

Worrall's reply (2010, p. 65): at least, the confirmation of H by the set of all evidences used by scientists is not person-relative (?)

(4) The UN criterion seems to be in conflict with probabilistic confirmation.

In what follows I present a probabilistic account of genuine confirmation that

– naturally entails the UN-criterion
– provides (better) solutions to the above objections, and
– can be understood as a plausible strengthening of Bayesian confirmation concept.

4. Grounding Use-Novelty on Epistemic Probability

Central probabilistic argument against ex-post fitting:

In a context C_1 where Hx, in order to explain the actual evidence E, is strengthened to Hc by fitting to E itself, Hx cannot increase E's probability

⇒ *because* the general theory Hx can be fitted to every possible evidence $E_1,\ldots,E_n$ (obtainable from a given test/experiment t) that H intends to explain;

hence $P(E_i \mid Hx) = P(E_i)$ (whence $P(Hx \mid E_i) = P(Hx)$ *

Relation to Mayo's account: even if Hx is false, it can be fitted to E_i (via Hc).

In contrast, in a context C_2 where Hx, in order to explain E, was fitted to another evidence E* obtained in a different test t*, the resulting Hx-strengthening Hc* *cannot* fit every possible outcome of t, because c* has already been fixed.

hence if Hc*fits E, Hx is highly confirmed by E.

Relation to Mayo's account: if Hx is false, it is highly improbable that Hc* fits E.

* where in general, $P(E_i \mid Hx, C) = P(E_i \mid Hc_C)$

Crucial point that solves objection (3) concerning person-relativity:

P is NOT a semantic but an epistemic relation between propositions

P depends on an epistemic background context C that determines how the evidence was obtained (by random or artificial selection), which role the evidence played in the construction of the hypothesis (and …)

P above is relativized to C – we can either write P_C or $P(\ldots \mid \ldots \wedge C)$.

$P(Hx \mid E \wedge C_1)$ = low; $P(Hx \mid E \wedge C_2)$ =high, where

C_1: E is explained by a Hx-specialization obtained by fitting to E itself

C_2: E is explained by a Hx-specialization obtained by fitting to independent E*.

Standard Bayesian objection: nevertheless, independently from the status of Hx:

the fitted hypotheses Hc makes E highly probable,

so $P(Hc| E) > P(Hc)$ *by probability calculus*; i.e. E Bayes-confirms Hc.

My counter-argument: this kind of confirmation rest on mere *content-cutting:*

E confirms Hc because E confirms that content part of Hc which it is itself!

But: genuine confirmation is *content-transcending* → it produces probability-transfer from E to those parts of Hc which go beyond E.

Application to the fitting-problem: the essential content part of Hc that goes beyond E is Hc; and the probability of this content-part is *not* increased by E.

5. Genuine Confirmation

Definition of (full) genuine confirmation: E confirms H (fully) genuinely iff E increases the probability of all those contingent *content parts* of H that go beyond E (are not logically contained in E).

Important: Content parts are not arbitrary logical consequences but *relevant elementary* consequences in the sense of Schurz (1991, 2010) (similar Gemes 1993).

$\rightarrow$ otherwise this definition falls prey to the *Popper-Miller-objection* (Miller 1990)

if H entails E, then $H \leftrightarrow E\wedge(\neg E\vee H)$, and $P(\neg E\vee H \mid E) < P(\neg E\vee H)$.

S is a content part of H iff

(a) H entails S,

(b) no predicate (includ. prop. variables) in S is replaceable on some of its occurrences by an arbitrary other predicate (of same degree) salva validitate, and

c) S is not logically equivalent with a conjunction of sentences shorter than S

E.g.: $\{p\vee q, p\vee\neg q\}$ are *not* content parts of p – the only content part is p

ContParts($\{p\rightarrow q, q\rightarrow r\}$) = $\{\neg p\vee q, \neg q\vee r, \neg p\vee r\}$

ContParts($\{p\vee\neg q, p\vee q\}$) = $\{p\}$

In propositional logic content parts coincide with (relevant) clauses.

6. Applications

6.1 Irrelevant conjuncts (tacking by conjunction, Glymour 1981):

Let E = grass is green and X = an absurd theory, e.g. the doctrine of Jehova's witnesses.

Then the hypothesis H := E $\wedge$ X is Bayes-confirmed by E.

$\rightarrow$ *no genuine confirmation*, because E does not increase X's probability (mere 'content-cutting').

Note: decomposition of H into content-part-conjunction E$\wedge$X is not always possible – but at least, H will always have *some* content parts not entailed by E.

Iteration: Often, E genuinely confirms only a conjunctive part of theory H$\wedge$X.

E.g. H = "Combustion involves oxidation" X = "Phlogiston exists"

We want to say, E(chemical evidence) is a partial genuine confirmation of H$\wedge$X .

Note: This is also a problem for Bayesians (cf. Crupi/Tentori 2010 vs. Fitelson 2002).

Definition of partial genuine confirmation: E confirms H partially genuinely iff H has at least some contingent content part that is fully genuinely confirmed by E.

6.2 Curve Fitting […already discussed, very brief:]

This case has been already discussed – e.g., recall:

In context C_1 in which H_{lin} was obtained by fitting Lin to E, H_{lin} is not genuinely confirmed by E because its the probability of its content part Lin's is not increased by E. (Likewise for H_{pol} …).

6.3 Speculations versus and Scientific Theories

Recall application to creationism [already discussed]:

Hx = God makes it that x   Hc = God makes it that E   (and …)

Hx would only be confirmable by E if E were a *use-novel* prediction of some specialization of Hx (obtained from fitting to independent evidence E*).

⇒ but refined creationism doesn't make any use-novel predictions

so it is not genuinely confirmable   (I think, is the right demarcation criterion).

If a strengthened version of creationism *would* make independently testable predictions, it would belong to the family of empirically falsifiable creationisms.

Challenge: Versions of creationism that entail inductive generalizations entail predictions that go beyond E. Aren't they therefore partially genuinely confirmable?

Example:	E:	So far the sun was rising every day
	is explained by H:	God makes it that the sun rises every day
	entails prediction E':	The sun will rise also in the future,

We shouldn't treat this as genuine confirmation, because …

Necessary strengthening for theoretical hypotheses:
A *theoretical* hypothesis H (that contains theoretical concepts) is partially genuinely confirmed by E iff H contains at least some contingent content part that is fully genuinely confirmed by E *and is not inferable from E by mere inductive generalization*.

⇒ otherwise, the same explanation would be possible without the introduction of a theoretical concept (Ockham's razor).

⇒ So: theoretical hypotheses are only genuinely confirmable by evidence E that is *qualitatively novel* in regard to fitting evidence E' (not inductively inferable from E').

Example of qual.	E:	So far the sun was rising every day
novel prediction:	is explained by H:	The earth rotates around itself
	entails e.g. E':	All stars turn over the nightly horizon with equal rotating speed.

Here, H is genuinely confirmed by E'.

6.4 Inductive generalization – a solution to objection (2):

Why is fitting legitimate in the inductive inference from the (relative) frequency of a *sample* to the frequency (limit) in the total *domain*?

The general theory Hx says here: the frequency of the respective property F in the domain has *some* value x.
The special hypotheses Hc is obtained by fitting x to E's sample frequency (k(n).

Also in this case, the assertion Hx is not confirmed by a particular sample E – *but:*
⇒ if the domain is *finite*, Hx is a logical truth and, thus, not a contingent content part of Hc.
So H[c] is genuinely confirmed by E (because every contingent content part of Hc going beyond E is inductively confirmed by E.

Note: in this case, Worrall's implication Hx→Hc is logically equivalent with Hx (this would have been a possible reply of Worrall to Mayo)

⇒ If the domain is *infinite*, Hx asserts that the F-frequency in random sequences taken from the domain converges to a limit
Here, Hx has to be assumed in the background context
Then Hx is no longer contingent, though it isn't a logical truth.
Note: Hx is confirmable by different means, e.g. by checking convergence rates in finite random sequences.

Thank you!

References

Behe, M. (1996): *Darwin's Black Box*, Free Press, New York.

Crupi, V., and Tentori, K. (2010): "Irrelevant Conjunction: Statement and Solution of a New Paradox", *Philosophy of Science* 77, 1-13.

Dembski, W. (1998): *The Design Inference*, Cambridge Univ. Press, Cambridge.

Fitelson, B. (2002): "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction", *Philosophy of Science* 69, 611-622.

Gemes, K. (1993): "Hypothetico-Deductivism, Content, and the Natural Axiomatization of Theories", *Philosophy of Science*, 54, 477-487.

Glymour, C. (1981): *Theory and Evidence*, Princeton Univ. Press, Princeton.

Hitchcock C. and Sober, E. (2004): "Prediction Versus Accommodation and the Risk of Overfitting", *British Journal for the Philosophy of Science* 55: 1-34.

Howson, C. (1990): "Fitting Theory to the Facts: Probably not Such a Bad Idea After All", in: C. Savage /ed.), *Scientific Theories*, Univ. Minnesota Press, Minneapolis.

Howson, C. und Urbach, P. (1996): *Scientific Reasoning: The Bayesian Approach*, Open Court, Chicago (2. Aufl.).

Ladyman, J., and Ross, D. (2007): *Every Thing Must Go. Metaphysics Naturalized*, Oxford University Press, Oxford. (With D. Spurrett and J. Collier.)

Lakatos, I (1977): "Science and Pseudoscience", in: derselbe, *Philosophical Papers Vol. 1*, Cambridge Univ. Press, Cambridge, 1-7.

Mayo, D. (1996): *Error and the Growth of Experimental Knowledge*, University of Chicago Press, Chicago.

Mosier, C. I. (1951): "Problems and Designs of Cross-Validation", *Educational and Psychological Measurement* 11, 5-11.

Musgrave, A. E. (1974): "Logical versus Historical Theories of Confirmation", *British Journal for the Philosophy of Science* 25, 1-23.

Paulßen, J. (2010): *Das Problem der Kurvenanpassung*, PhD Univ. Duesseldorf (http://docserv.uni-duesseldorf.de/servlets/DerivateServlet/Derivate-16750/diss.pdf)

Pennock, R.T. (2011). "Can't Philosophers Tell the Difference between Science and Religion? Demarcation Revisited", *Synthese* 178, 177-206.

Schurz, G. (1991): "Relevant Deduction", *Erkenntnis*, 35, 391 - 437.

Schurz, G., & Weingartner, P. (2010): "Zwart and Franssen's Impossibility Theorem …", *Synthese*, 172, 415-436.

Swinburne, R. (1979): *The Existence of God*, Clarendon Press, Oxford

Stegmüller, W. (1983), *Probleme und Resultate …. Band I: Erklärung - Begründung - Kausalität* (2. verbesserte und erweiterte Auflage), Berlin: Springer.

Unwin, S. T. (2005): *Die Wahrscheinlichkeit der Existenz Gottes. Mit einer einfachen Formel auf der Spur der letzten Wahrheit*, discorsi, Hamburg.

Worrall, J. (2010): "Theory-Confirmation and History", in: C. Cheyne and J. Worrall (eds.), *Rationality and Reality*, Springer, New York, 31-61.