**Titel / Title:**

# Genuine Confirmation and the Use-Novelty Criterion

**Autor / Author:**

Gerhard Schurz

TPD PREPRINTS

Annual 2011 No.2

Edited by Gerhard Schurz and Ioannis Votsis

Vorveröffentlichungsreihe des Lehrstuhls für
Theoretische Philosophie an der Universität Düsseldorf
Prepublication Series of the Chair of
Theoretical Philosophy at the University of Düsseldorf

# Genuine Confirmation and the Use-Novelty Criterion

Gerhard Schurz (Universität Düsseldorf)

*Abstract*
According to the Bayesian concept of confirmation, rationalized versions of creationism come out as empirically confirmed. From a scientific viewpoint, however, they are pseudo-explanations because with help of them all kinds of experiences are explainable in an *ex-post*. fashion, by way of ad-hoc fitting of an empirically empty theoretical framework to the given evidence. An alternative concept of confirmation that attempts to capture this intuition is the use novelty (UN) criterion of confirmation. Serious objections have been raised against this criterion. In this paper I suggest solutions to these objections. Based on them, I develop an account of *genuine* confirmation that unifies the UN-criterion and Mayo's severe-test criterion with a refined probabilistic concept of confirmation that is explicated in terms of the confirmation of evidence-transcending content parts of the hypothesis.

## *1. The Problem: Bayesian Confirmation of Irrational Beliefs*

Neo-creationists have applied Bayesian confirmation methods to confirm *refined* versions of *creationism*. With help of Bayes' formula Unwin (2005) has calculated the posterior probability of God's existence as 67%. Swinburne (1979, ch. 13) is more cautious; his major argument is based in the claim that certain experiences *increases* the probability of God's existence.  Can something like this really count as a serious confirmation? To answer this question we first distinguish two kinds of creationisms:

*Empirically criticizable* creationisms are testable at hand of empirical consequences. Literal interpretations of the genesis and other holy scriptures make plenty of empirically testable assertions (e.g. concerning the age of the cosmos or the creation of biological species), but more-or-less all of them have been scientifically refuted. These empirically criticize creationisms don't constitute a problem for Bayesian (and other) confirmation accounts.

We are dealing here with *empirically uncriticizable* creationisms. These are *ratio-*

*nalized* versions of creationism that carefully avoid any conflict with established empirical knowledge, but nevertheless entail empirical consequences. How is this possible? Take the empirically vacuous creator-hypothesis and enrich it with scientifically established empirical facts as follows:

(1) *Hypothesis of rationalized creationism:* God has created our world with the following properties (E): … [here follows a list of as many scientifically established facts as possible].

History of rationalized theology is full of pseudo-explanations of that sort. Contemporary neo-creationists proudly announce that they can ex-post explain even so difficult facts as the fine-tuning of the constants of nature (Dembski 2003), or the intricate complexity of the human eye (Behe 2003). The most 'advanced' rationalized creationists have stipulated a God that creates the living beings by the mechanism of Darwinian evolution (Isaak 2002). Intuitively we feel that something is wrong with these kind of ex-post 'explanations', but what could it be? The problem is that even from the viewpoint of one of the most influential confirmation theories in philosophy of science, namely Bayesian confirmation theory, creationist pseudo-explanation comes out as being confirmed by the evidence which it 'explains'. To see why, we need some technicalities: in what follows H (or $H_1$, $H_i$…) stand for hypotheses, E ($E_i$,…) for empirical evidences, P(H) for H's prior probability, P(H|E) for H's posterior probability given E, and P(E|H) for E's probability given H – the so-called likelihood. Sometimes, but not always, this likelihood objectively determined. For example, the likelihood of E = "throwing heads", given H = "throwing a regular coin" is 1/2 by the laws of statistics. In particular, the likelihood of E given a hypothesis H which logically implies E is always 1. A proposition A is called *epistemically contingent* iff $0 < P(A) < 1$, i.e. its probability is different from 0 and 1.

There exist two different (basic) Bayesian confirmation concepts: H is *absolutely* confirmed by E iff P(H|E) is sufficiently high (at least higher than 1/2), where this conditional probability is computed by the famous Bayes-formula as follows:

(2) *Bayes-formula:* $P(H|E) = P(E|H) \cdot P(H) / P(E)$,

where $P(E) = \Sigma_{1\leq i\leq n} P(E|H_i)\cdot P(H_i)$.

Thereby, $\{H_1,\ldots,H_n\}$ is a partition of alternative hypotheses containing H (i.e., $H = H_k$ for some k, $1\leq k\leq n$).[1] $P(H_i)$ is so-called the prior probability of $H_i$. It is widely accepted that these prior probabilities are the most problematic part of Bayesian confirmation theory, because degrees of belief which are 'prior to experience' are subjective and merely reflect one's own prejudices.

The second Bayesian confirmation concept is the *comparative* one. According to this comparative concept H is confirmed by E iff E increases H's probability, i.e. iff $P(H|E) > P(H)$. The confirmation concept has the advantage that it is independent from the (subjective) choice of H's prior probability; it depends only on the likelihoods. For the Bayes-formula in (2) entails that if H and E are epistemically contingent, then $P(H|E) > P(H)$ holds iff $P(E|H) > P(E)$. Moreover, $P(E|H) > P(E)$ is provably equivalent with $P(E|H) > P(E|\neg H)$ (and likewise, $P(H|E) > P(H)$ is equivalent with $P(H|E) > P(H|\neg E)$).[2] On this reason, most contemporary Bayesians prefer the comparative concept of confirmation, or quantitative refinements of it.[3] When speaking of "Bayesian confirmation" in what follows we always mean this comparative concept.

However, the comparative Bayesian confirmation concept has the following awkward consequence which allows all sorts of Bayesian pseudo-confirmations:

---

1 That $\{H_1,\ldots,H_n\}$ is a *partition* means that the $H_i$ are pairwise incompatible and jointly exhaustive, relative to a (possibly empty) backgorund knowledge on which P is conditionalized.

2 This follows from $P(E) = P(E|H)\cdot P(H) + P(E|\neg H)\cdot P(\neg H)$.

3 Two quantitative refinements of comparative confirmation are the *difference* measure $P(H|E)-P(H)$.and the *ratio* measure $P(H|E)/P(H)$.

(3) *Bayesian pseudo-confirmation:* Every epistemically contingent hypothesis H which entails an epistemically contingent evidence E is confirmed by E in the comparative-Bayesian sense.

(3) is an easy consequence of (2), because if H $\Vdash$ E (H entails E), then $P(E|H) = 1$, whence $P(H|E) = P(H)/P(E) > P(H)$ because $P(E) < 1$. This consequence can be exploited by proponents of all sorts of rational speculation. Every hypothesis, be it as weird as you want, will be confirmed by a given evidence E if it only entails E (cf. Schurz 2008, §7.1). For example, the fact that grass is green confirms the hypothesis that God wanted this and had brought it about that grass is green. Of course, the same fact also confirms the hypothesis that not God but a *flying spaghetti monster*[4] has brought it about that grass is green – or that a god together spaghetti-monster have made grass green, and so on … until the scientific explanation of the green colour of grass in terms of chlorophyll. All these explanatory hypotheses $H_i$ get comparatively confirmed by E. If they have a different conditional degree of belief $P(H_i|E)$, then (according to the Bayes-formula (2)) this can only be because they have different prior probabilities, since the likelihood $P(E|H_i)$ is 1 for all of them, and the value of P(E) is independent from the chosen hypothesis.

Bayesian philosophers of science are aware of this strange result. They usually reply that scientific hypotheses have a significantly higher prior probability than religious hypotheses (cf. Howson and Urbach 1996, 141f; Sober 1993, 31f). This reply is, however, doubly questionable:

(1.) Prior probabilities are subjective; and it seems to be inappropriate to ground the distinction between scientific hypotheses and speculations on subjective prejudices. From the religious point of view creationism has a higher prior probability than

[4] The *church of the flying spaghetti-monster* is a movement initiated by a physicist who intended to turn creationist teaching requirements into absurdity. See *www.venganza.org/aboutr/open-letter.*

evolution theory.[5]

(2.) Independent from objection (1.), it seems that rationalized creationism is not just "a little bit less" confirmed than evolution theory. Rather, it is *not confirmed at all* by way of these ex-post explanation, like any other of the absurd hypotheses mentioned above.

Let me add two points: *First*, the same shattering result does not only undermine Bayesian confirmation but also the (naive) hypothetico-deductive (hd) confirmation criterion. According to the latter, E confirms H if H entails E, provided H is not a contradiction and E not a tautology. Observe that (naive) hd-confirmation follows from (naive) Bayesian confirmation as a special case. *Second*, pseudo-confirmation is also a problem when H does not entail E but makes it only highly probably. For if H is epistemically contingent then E confirms H as soon as $P(E|H) > P(E)$. So assuming the prior probability of the fact that Grass is green is not too high, then this fact does also confirm probabilistic weakenings of the above pseudo-explanations, such as "a spaghetti-monster whose wishes become reality in 99% of all cases has wanted that grass is green" (etc.).

Of course, all that doesn't refute the moderate claim that the Bayesian confirmation criterion isn't at least a *necessary* condition for genuine conformation; what it shows is that the Bayesian criterion is *not sufficient*. We conclude that the Bayesian confirmation criterion is too weak to demarcate genuine confirmation from pseudo-confirmation.

A final remark on the *demarcation* problem – i.e., on philosophical attempts to demarcate science from pseudo-science. It is well known that in the last two decades – in part because of intrinsic difficulties like the above – work on the demarcation problem has become out fashioned in philosophy of science. More recently, however,

[5] For example, when Unwin (2005) computed the posterior probability of God's existence to be 67%, he (naively) assumed a 1:1 priori probability of Gods existence. This motivated the editor of the magazine *Sekptic*, Michael Shermer, to set up a counter-computation with different priors that resulted in a posterior probability of God's existence of merely 2%

creationists have relied on anti-demarcationist philosophers of science such as Laudan (e.g. 1983) to support their creationist teaching demands (cf. Pennock 2011, 180). This quite embarrassing fact provoked a new wave of discussion on adequate criteria of demarcation. In this new wave, demarcation plays a different role than for the neo-positivists in the early 20th century: not as an inner-philosophical weapon against metaphysics, but as a social guideline concerning the content of teaching curricula. This paper, though primarily on confirmation, should also be understood as a contribution to this debate.

## *2. Alternative Confirmation Concepts: Novel Predictions und Use Novelty*

It is the fundamental characteristic of pseudo-explanations like that of rationalized creationism that they are purely *ex-post*, resulting from fittings of empty speculations to already known data. Such ex-post explanations could never figure as predictions. On this reason, many philosophers of science have suggested to regard the failure of delivering predictions as the discrimination mark between pseudo-confirmation versus genuine confirmation.[6] According to the criterion of *novel predictions,* in short: the *NP*-criterion, a confirming evidence for a hypothesis H must be a novel prediction of H, i.e. a true empirical consequence of H which was unknown at the time point when H was introduced or constructed.

Indeed, rationalized creationism cannot predict anything because it doesn't tell us anything about the properties of the creator except that he has causes the facts asserted in the explanandum. Therefore the creationist hypotheses "God wanted E " can only be given in retrospect, when E is already known. Note that on the same reason, creationistic ex-post explanation don't provide genuine explanatory unification, but only *pseudo-unification*: since for every new empirical fact a new wish of God has to be assumed, God's wishes can never be more unified, or better understood, than the

[6] Cf. Musgrave (1974), who cites Descartes, Leibniz, Whewell, Duhem; more recently Watkins (1964), Lakatos (1977) and Ladyman and Ross (2007, §2.1.3).

brute empirical facts themselves. This note is important because often creationists wrongly assert that 'explanations' by God's will are simpler or more unified than scientific explanations (cf. Swinburne 1979, ch. 14; Dembski 2003). But this is wrong – as we shall see, the opposite is the case.

Keep in mind that in the NP-criterion the notion of prediction is not understood in the temporal but in the *epistemic* sense: it is not required that the inferred evidence E refers to the future, but merely that the hypothesis H has been known already *before* E was known, and E was inferred from H afterwards – this is sufficient to exclude ex-post fitting. Predictions in the epistemic include not only temporal predictions, but also *retrodictions*, i.e. inference concerning the past, which are yielded in great wealth by evolution theory.

The NP-criterion has to face several objections. One objection argues that the confirmation relation must be independent from *pragmatic* aspects such as the time point the evidence has first been recognized, because confirmation is a purely *semantic* relation between the propositional contents of the hypothesis H and the evidence E (and possibly background beliefs B). This objection – though correct in a certain respect – is too general to be true. Some non-semantic – and in this sense 'pragmatic' – factors *do* play an obvious role for confirmation: factors concerning dependencies between the hypothesis and the way the data had been collected. For example, the confirmation of a statistical hypothesis H depends not only on the sample frequency E, but also on whether the sample was a *random* sample (as opposed to a selected sample of H-favourable instances). This does not mean that the semantic framework of probabilistic confirmation theory can no longer be applied, but merely, that the algebra of propositions over which the probability function is constructed has to include propositions describing these procedural facts. For example, the (epistemic) probability that a given coin is fair (H), conditional on the evidence E(a) that in a given sequence (a) of 100 given coin tosses 50 have been heads, depends on the additional procedural information R(a) that the sequence (a) was randomly selected. Thus we have $P(H|E(a) \wedge R(a))$ = high, while $P(H|E(a) \wedge \neg R(a))$ = low. In this way, probabilism is restored, but without its narrow semantic clothing.

In one respect, however, the previous objection is right: the *mere time point* of when an empirical fact has first been recognized is *per se* irrelevant for its confirmational value; what matters is whether this fact has been *used* in the construction of the hypothesis or not. There exist indubitable examples of confirmations of scientific theories by facts which the inventors of the theory did not use (and often didn't even know) when constructing their theory, although these facts were known long before. Examples of this sort are the confirmation of special relativity theory by the Michelson-Morley experiment, or of general relativity theory by Mercury's abnormal planetary orbits (cf. Musgrave 1974, 11f).

Therefore Worrall (2010a) has suggested the criterion of *use novelty*, in short the UN-criterion, as an improvement of the NP-criterion (the idea goes back to Zahar 1973). According to the UN-criterion an evidence E can only confirm a hypothesis H if E has not been used in the construction of H (in short, if E is 'use-novel'). The UN-criterion is a clear improvement of the NP-criterion: UN is weaker than NP (temporal novelty of E implies E's use-novelty, but not vice versa), but UN still excludes what is at stake, namely ex-post fitting of a hypothesis to given data. Moreover, the UN-criterion is equivalent to two formulations of the same idea that are likewise attractive, namely:

(a) *independent testability*: an experiment deciding E versus non-E (where P(E|H) is assumed to be high) is said to be an independent test of H iff E is use-novel w.r.t. H, and

(b) *potential predictiveness*: E is use-novel w.r.t. H iff E *could* have figured as a prediction from H.

## *3. Objections against the UN-criterion and Worrall's Account of Parameter-Adjustment*

Also against the UN-criterion serious objections have been raised. Howson (1990) has pointed out that the UN-criterion is violated by hypotheses arising from inductive generalizations: here the frequency hypothesis about the entire domain (the 'population') is obtained by adjusting the unknown population-frequency parameter to the observed sample frequency. This seems to be a perfect example of ex-post fitting, and yet the general hypothesis is doubtlessly confirmed by the sample frequency (provided one believes in induction, i.e., one assumes a suitably inductive probability measure; see below). In another counterexample that is due to Mayo (1991, 534f), a hypothesis about the average SAT (student admission test) score of her class is tested by adding the SAT student's scores and dividing through the number of students – in this example, the hypotheses is logically inferred from the observation result; a drastic violation of use-novelty, and yet an example of perfect confirmation, namely logical demonstration.

In my view, the defence that is given by Worrall to these objections is not convincing. Mayo's example, so Worrall (2010, 69f; 2010b, 134), is a case of "demonstration", not "test" or "confirmation". For me this looks like a purely linguistic manoeuvre: why should entailment of H by E not be considered as the extreme case of a confirmation of H by E with P(H|E) = 1? Concerning Howson's example, Worrall acknowledges that here the hypothesis is not really deduced from the evidence. But, so Worrall (2010b, 132f), it is "quasi-deduced" from it; so also here we are not facing a genuine situation of "test" or "confirmation". It seems to me an entirely untenable step to subsume inductive generalization procedures under the unclear rubrique of "quasi-deduction" – apart from the fact that even if this step were successful, it were a merely linguistic manoeuvre – nothing would be gained by it, because Howson's example is an indubitable a case of genuine (inductive) confirmation.

So let us ask: what is really going wrong with the UN-criterion in the examples of Howson and Mayo? In section 5 I will suggest two improvements of Worrall's account, which give a preliminary answer to this question. These improvements are based on Worrall's account of theory-construction by *data-driven parameter-adjustment*. According to this account, the intended applications of the UN-criterion are cases in which a general background hypothesis or theory contains one or several freely variable parameters which can be adjusted ex-post towards given evidence E, in such a way that E is made highly probable or even follows from the parameter-adjusted theory. In what follows we let Tx stand for the *general* (background) theory or theory-frame with variable parameters x, E for the evidence, and Tc for the parameter-adjusted *specialization* of Tx with constant parameters c. To indicate that x was obtained from fitting Tx onto E we also write $c_E$ for c and $Tc_E$ for Tc (recall, $Tc_E$ makes E highly probable). We note the following logical facts:

1.) Tx abbreviates $\exists xT(x)$; i.e., we assume the general theory existentially quantifies over the variable parameter. Hence Tx follows from Tc.

2.) There may be several such parameters (i.e., x stands short for $x_1,x_2,\ldots$, and likewise, c for $c_1$, $c_2,\ldots$).. and they may be of 1st or higher order, ranging over numerals, individuals or predicates.

We illustrate this at hand of two examples. In the case of *creationism*, thegeneral theory Tx asserts (in its simplest version) that there God created the world with various variable (unknown) facts X ($\exists$X: God created X), and the parameter-adjusted theory $Tc_E$ says that God created the world with some specific (known) facts E (or $E_1 \wedge E_2 \wedge \ldots$). With "creation" we mean, of course, intentional creation, i.e. God brought it about that X because he wanted it that X. In the more concrete example of curve fitting (to be illustrated in the net section) the general theory Tx is an assertion of a certain *type* of functional (e.g. linear) dependency between two variables X and Y, and $Tc_E$ is the particular (e.g. linear) curve fitting optimally to the given set of data points E (modulo a random deviation).

Worrall (2010, 49ff) suggest the following "dualistic" confirmation theory: given $Tc_E$ results from Tx by ex-post parameter-adjustment, then:

(a) the special theory is confirmed by E only on the condition that the general background theory is true, in other words, E confirms the *implication* $Tc_E \rightarrow Tx$;

(b) but E does not and *cannot* confirm the general theory Tx.

For example, the fine-tuning of the constants of nature (E) confirms that God created this fine-tuning ($Tc_E$) only on the presupposition that God created the world (Tx), but it does not confirm that God created the world. In other words, the confirmation that E provides for $Tc_E$ conditionally on Tx does not spread to the background theory Tx. Transport of confirmation from $Tc_E$ to Tx can only be achieved by *independent* evidence E' that is entailed (or made probable) by $Tc_E$ and was not used in for parameter-adjustment: if this happens, then $Tc_E$ as well as the background theory Tx is unconditionally confirmed by E', i.e., the confirmation flows from $Tc_E$ to Tx (Worrall 2010a, 53f). Such independent evidence E' does (usually) not exist in the case of rationalized creationism. But it exists in the case of curve-fitting – the paradigm case of parameter adjustment in science, to which we turn now.

## *4. The Example of Curve Fitting*

Curve fitting is usually done with polynomial functions, on the reason that polynomial function can finitely approximate arbitrary functions with arbitrary precision. A polynomial function of *degree n* in two real-valued variables X,Y has the general form $Y = c_0 + c_1 \cdot X + c_2 \cdot X^2 + \ldots c_n \cdot X^n$ (for n=1 the polynomial is linear, for n=2 quadratic, etc.). The variables X and Y (so-called 'random variables) are themselves functions (physical magnitudes) over the individuals $d_1$, $d_2$,… of a given domain D, who are the bearers of the physical magnitudes (e.g., measurement results). So the formula "Y = f(X)" is just a shorthand for $\forall d \in D(Y(d) = f(X(d))$. Let us assume that the variables X and Y are related by some true but unknown polynomial function $f^n: X \rightarrow Y$ of unknown degree n, plus some unknown Gaussian *random* (error) dispersion $\sigma$ around this function; i.e. $Y = f^n(X) + r(\sigma)$, where $r(\sigma)$ is a Gaussian distributed random term with mean 0 and standard deviation $\sigma$. Assume we have measured m data points in the X-Y-coordinate system, i.e. our evidence is $E = (X_i, Y_i):1 \leq i \leq m)$. Which (polynomial) X-Y-curve should be inductively infer from these data? It is a well-known fact that every set of m data points can be approximated by every polynomial function of variable degree up to variable remainder dispersion $\sigma$ which is the smaller, the higher the degree of the polynomial, and becomes zero if n ≥ m+1, i.e., if the polynomial has at least as many freely variable parameters as there are data points. Consider the set of data points and the two curves in fig. 1.

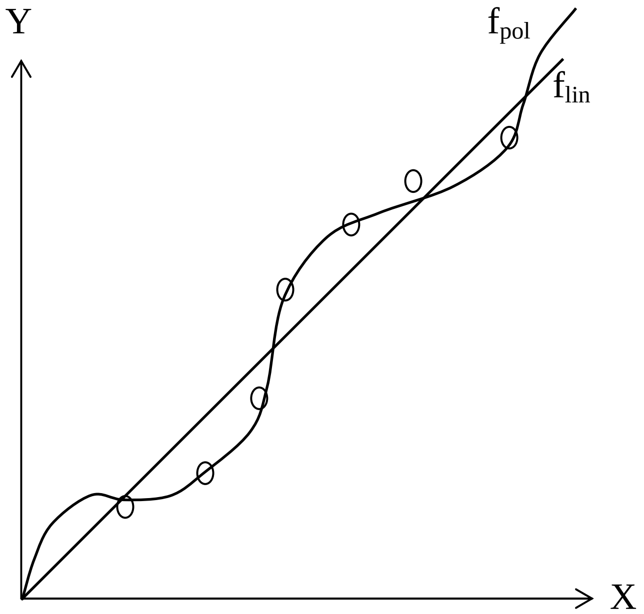


*Fig. 1: Linear vs. (high-degree) polynomial curve fitting*

Both curves result from fitting the free parameters of the respective polynomial (linear vs. high-degree) optimally to the data, i.e., from adjusting them in a way which minimizes the so-called *SSD*, that is the sum of squared deviations of the data points from the curve ($\Sigma_{1\leq i\leq m}Y_i – f(X_i))^2$). Of course, the high-degree polynomial curve $f_{pol}$ approximates the data better than $f_{lin}$, because it had more free parameters to be adjusted – but is therefore $f_{pol}$ therefore better confirmed than $f_{lin}$? NO, because $c_{pol}$ may have *overfitted* the data, that is, it may have fitted on *random accidentalities* of the sample instead on the systematic dependency between X and Y (cf. Hitchcock and Sober 2003). Generally speaking, the method of SSD-minimizing gives one the polynomial $f^n$ with the highest data-likelihood ($P(E|f^n)$) among all curves of the same *type* (i.e. degree); but it cannot tell us which is the right type of curve (cf. Glymour 1981, 322). Without independent information about the true dispersion (which is not available in our setting) nothing about a curve's confirmation value is inferable from the achieved degree of approximation. This is only possible by testing the fitted curve at hand of a *new* data set $E_2$ that was not used for parameter-adjustment. This is show in fig. 2 – the new data points are drawn in grey, the old ones in white:

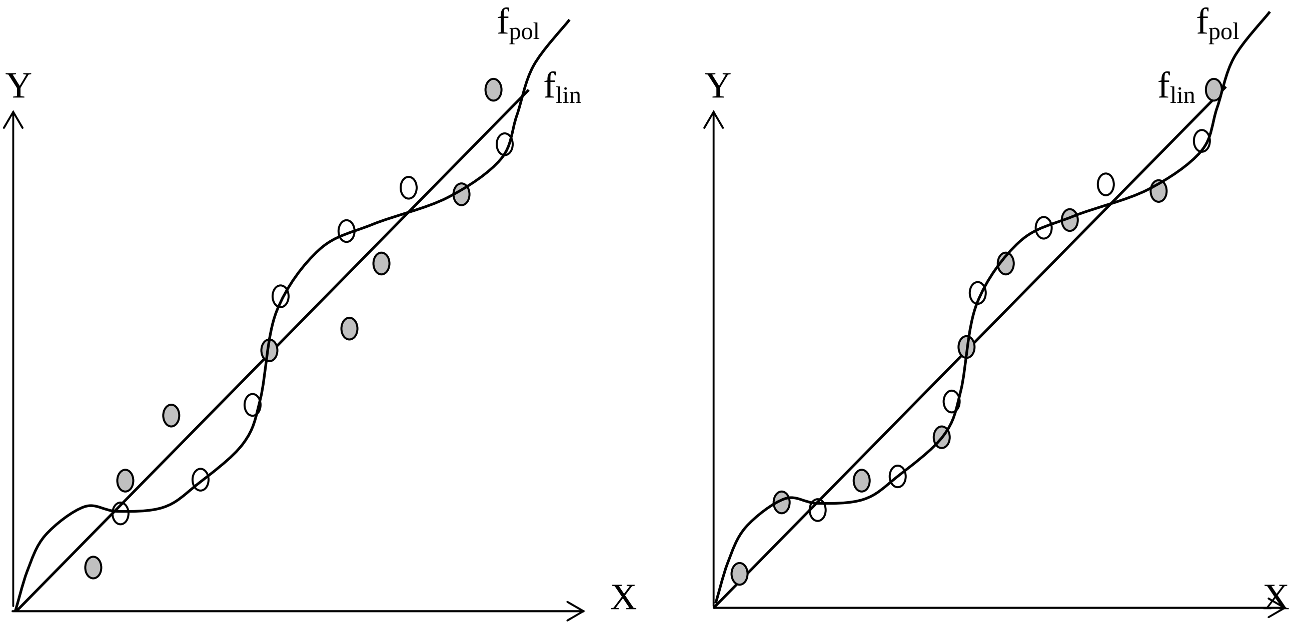


*Fig. 2:* New data ($E_2$) in grey, old data ($E_1$)in white.
*(2a):* $f_{lin}$ and 'Lin' is confirmed by $E_2$ *(2b):* $f_{pol}$ and 'Pol' is confirmed by $E_2$

In the left case (2a), the new data constitute independent confirmation of the linear curve ($f_{lin}$) and the underlying general linearity hypothesis ('Lin'), because they lie far off the wriggled line of the polynomial curve, but are within the standard deviation expected by $f_{lin}$. In the right case (2b), the new data lie quite tightly on this wriggled line and, hence, provide independent support for the polynomial curve.

The method of confirming fitted curves at hand of independent data sets is an excellent example of an application of the UN-criterion to parameter-adjustments in science. We now reconstruct this example within our general framework. Let $T_{lin}c_1$ stand for the special linearity hypothesis or 'theory' (abbreviated as $f_{lin}$ in fig. 1); so $T_{lin}c_1$ has the form $Y = c_1 \cdot X + c_0 + r(s)$, where $c_1$, $c_0$ and s are the constants of the optimal linear curve in regard to the data set $E_1$ (as explained). The underlying general theory is denoted by $T_{lin}x_1$ ('Lin' in fig. 2) and asserts that the X-Y-dependence is linear with unknown random deviation; it is obtained from $T_{lin}c_1$ by existential quantification over the parameters and has the form $\exists\alpha_o\exists\alpha_1\exists\sigma[Y = \alpha_1\cdot X+\alpha_0+r(\sigma)]$. Likewise, the special high-degree polynomial curve $f_{pol}$ is denoted as $T_{pol}c_2$ and has the form $Y = \alpha_k\cdot X^k\ldots+\alpha_1\cdot X+ \alpha_0+r(s)$ (for a given k); and the underlying general theory $T_{pol}x_2$ has the form $\exists\alpha_o\ldots\exists\alpha_k\exists\sigma\ [Y = \alpha_k\cdot X^k\ldots+\alpha_1\cdot X + \alpha_0+r(\sigma)]$.

Without any restriction on the size of the obtained remainder dispersion s, both $T_{lin}x_1$ and $T_{pol}x_2$ may be fitted to *every* given data set E – at least if the criterion of approximation is given the method of SSD-minimization. Therefore the adjustment of a general hypothesis – asserting that the true degree is k – cannot be confirmed by adjustment to a given data set E. What only can be said, in accordance with Worrall's "dualistic" account, is that *if* the true curve is a polynomial of degree k ($H_kx_k$), *then* the polynomial of degree k with the highest likelihood in regard to data set E has this-and-this parameters and this remainder dispersion ($H_kc_k$), i.e., $T_kx_k\rightarrow T_kc_k$. Only if the fitted curve $T_kc_k$ is independently confirmed by a second data set E*, then $T_kc_k$ and the underlying general hypothesis $T_kx_k$ is unconditionally confirmed.

## *5. Two improvements of Worrall's Account*

It is not clear whether really all cases of ex-post theory-construction can be regarded as specialization of a given background-theory by parameter adjustment. In the following, I don't discuss this question, but restrict my approach to applications where this reconstruction is appropriate. Based on the achieved insight, I now turn to two suggested improvements of Worrall's account. Each improvement consists in *removing* certain claims of Worrall that I think are untenable, and in retaining on *only* those parts that have a probabilistic justification.

*5.1 Improvement 1:* Worrall (2010a, 50, 66) asserts that the implication $Tx\rightarrow Tc_E$ is always *deduced* from E, but this not right in all cases. In the example of curve-fitting (section 4), the implication "if the X-Y-dependence is linear, then the linearity-constants are such-and-such" is not deductively but only *inductively* entailed by the data – for even if the linearity assumption is right, the data may have resulted from an unluckily misleading sample. More generally, Worrall's account of confirmation is too narrowly hypothetico-deductive (recall his notion of "quasi-deduction"); the confirmation account should cover all sorts of probabilistic confirmations, including hypothetic-deductive confirmation as a special case. Related to this point: although we agree with Worrall that the conditional confirmation of $Tc_E$ given Tx by E, and the unconditional confirmation of Tx by some independent E', are different *cases* of confirmation, we do not agree with his view (2010a, 66) that the two cases rest on two different *concepts* of confirmation: in section 6 we argue that both are covered by the same concept of genuine confirmation.

The objections of Howson and Mayo can be solved by the improved Worrall account without any need to exclude entailment or inductive parameter-estimation as genuine cases of confirmation. In Mayo's SAT-example, the parameter-variable hypotheses Tx asserts "there exists an average SAT-score of my students", which is logically true. So Tx is not in need of confirmation, and the implication $Tx \rightarrow Tc_E$ is logically equivalent with $Tc_E$ (in other words, confirmation conditional on a tautology equals unconditional confirmation). Here Worrall's own account yields that E confirms $Tc_E$, and I don't know why Worrall hasn't chosen this much easier route of defence. In Howson's example specialized theory Tc (domain frequency) is of course no longer entailed by the evidence E (sample frequency), but arises from E by an inductive generalization. Worrall seems to assume that this produces a problem for his account, but again this is not true. The general hypothesis Tx now states that "there exists a frequency or frequency-limit", of the respective property in the domain. If the underlying domain is finite, this is again a logical truth, and is treated like Mayo's SAT-example. If the domain is infinite, Tx is not logically true but asserts the existence of a frequency-limit (in random sequences of the domain). Also in this case, Worrall's account yields the right result: for the existence-assumption concerning a frequency-limit is presupposed, but cannot be confirmed by inductive frequency estimations – if it can be confirmed at all, then only by more complicated observation sequences concerning convergence rates.[7]

*5.2 Improvement 2:* The second suggestion of improvement arises from the central question: *Why* is the general background theory Tx *not* confirmed by E via $Tc_E$, when $Tc_E$ is a parameter-adjustment of Tx to E? Worrall gives two different answers to this question. Sometimes (e.g. 2010a, 44) he says (a) Tx is not confirmed by E already because of the mere fact that Tx has been fitted to E (yielding $Tc_E$) in an *ad-hoc* way. In other places (e.g. 2010b, 148) he gives the more specific reason that by way of parameter-adjustment Tx could have *been fitted* to every possible data $E_i$ (that may result from the type of considered experiment). I argue that only *(b) is right while (a) is wrong.*

That (b) is right has a straightforward probabilistic justification: if Tx can be made fit with all possible outcomes $E_1,\ldots,E_n$ of a considered experiment, then the assumption of Tx cannot increase the prior probability of any of these outcomes, i.e. $P(E_i|Tx) = P(E_i) = P(E_i|\neg Tx)$ must hold, and this implies by probability theory that $E_i$ cannot confirm Tx (in the ordinary Bayesian sense), i.e. $P(Tx|E_i) = P(Tx)$ must hold, for all $E_i$, $1 \le i \le n$. Note that this observation is not incompatible with the problematic fact that each $E_i$ Bayes-confirms $Tc_{Ei}$ (i.e. $P(Tc_{Ei}|E_i) > P(Tc_{Ei})$) – but I will argue in section 6 that this not *genuine* confirmation because $E_i$ does not increase the probability of any content-parts of $Tc_{Ei}$ that goes *beyond* $E_i$. One the other hand, the parameter-adjusted theory $Tc_{Ei}$ cannot be fitted to arbitrary independent data sets $E_j$ but only to very few ones, whence fit of $Tc_{Ei}$ to $E_j$ is a clear confirmation of $Tc_i$ and $Tx_i$ by $E_j$.

Argument (a) is wrong precisely in cases where argument (b) doesn't apply, i.e.,

---

[7] In Bayesian probability theory, the existence of a frequency-limit follows with probability 1 from the assumption that the probability measure is *exchangeable* (i.e. invariant w.r.t. permutation of individual constants; Earman 1992, p. 89; Carnap 1971, pp. 117ff). The exchangeability-axiom goes beyond the standard Kolmogorov probability axioms; it expresses a weak *inductive* assumption insofar it assumes that prior to all experience all individuals have the same probabilistic tendencies (cf. Earman 1992, p. 108).

where the general theory Tx can be adjusted only some but not all possible experimental outcomes, and the factual outcome E was among those that fit well with Tx. As a first example, consider again Howson's case of inductive frequency estimation over an infinite domain, where Tx is the assertion "the exists some frequency-limit x". Tx can be adjusted to every observed sample-frequency (E) producing Tc, e.g., "the frequency limit is 0.6". But now, instead of Tx consider the slightly less general background theory T*x that asserts that the domain-frequency lies between 0.2 and 0.8; and assume again, that Tc (= T*c) was obtained by adjustment of T*x's frequency-parameter to the observed sample frequency of 0.6. It would make little sense to say that now the found sample-frequency of 0.6 (E) doesn't confirm that the domain-frequency is approximately 0.6 (Tc), but only that *if* the domain-frequency lies between 0.2 and 0.8, *then* it is around 0.6 ($T^*x \rightarrow Tc$), equivalent with "either the domain-frequency is not between 0.2 and 0.8, or it is approximately 0.6". The reason why this diagnosis would now be nonsensical is precisely that T*x cannot be fitted to all sample frequencies, but only to those lying between 0.2 and 0.8: T*x increases the likelihood of the latter ones and is, thus, itself confirmed by E (as well as Tc).

Another case in point is Worrall's example of the *wave-theory of light* Tx with the free parameter (x) of the wave-length of a monochromatic light source, e.g. a sodium arc (2010a, 47ff). Applied to the two-slit experiment Tx predicts that the wave-length $\lambda$ equals $d \cdot L/(L^2+D^2)^{(1/2)}$, where d = the distance between the two slits, D = the distance between the two-slit screen and the observation screen, and L = the distance between the intensity peak (fringe) at the centre and the first peak on either side. Let $Tc_E$ be the parameter-adjusted theory asserting a particular wave-length of sodium arc as the result of a particular measurement E. Worrall argues that also in this case, E confirms $Tc_E$ conditional on the acceptance of Tx, but it does not confirm Tx, the wave-theory of light itself (as opposed to the corpuscular theory). I don' think that is true because already Tx alone has *some* empirical content: for example Tx predicts the emergence of an inference-pattern – concentric circles of high intensity separated by circles of zero-intensity on the observation screen – independently from the specific value of the function $d\cdot L/(L^2+D^2)^{(1/2)}$ that equals $\lambda$. Tx could not be adjusted to an observed light pattern whose intensity continuously degrades from the centre to the margin, as predicted by the corpuscular theory of light.

What is common to the two cases is that the special theory Tc is obtained by parameter-adjustment of a general theory Tx that *cannot* be fitted to all but only to some possible experimental outcomes; in such a case both Tx and Tc are confirmed by the evidence. The same observation applies to Worrall's strategy of defence against an objection of Musgrave. Musgrave (1974, 14) gave this argument against Zahar's earlier version of Worrall's account. He objected that different scientists may arrive at different routes to the same specialized theory: one may use evidence $E_1$ to adjust Tx to Tc and confirm Tc by independent evidence $E_2$, while the other scientist may use $E_2$ to adjust Tx to Tc and use $E_1$ as independent confirmation. Worrall replied that what one should say in regard to sets of evidences is this: Tc and Tx are confirmed by an *evidence set* $\{E_1,\ldots,E_n\}$ iff this set contains at least *some* independent evidences, i.e. evidences that have not been used to adjust Tx's parameters – and this relation holds objectively, independently from the route that scientists take in their parameter-adjustments. I regard Worrall's view as entirely correct, but this view also commits one to the insight (b) above. To see this, simply assume that E is the set or conjunction of all the evidences (e.g. data-points) $E_1,\ldots,E_n$, where the first k of them ($k < m$) have been used to adjust Tx's parameters. Then Tx cannot be fitted to all possible combinations of n evidences (experimental outcomes), but only to all combinations of the first k of them, whence again, not only Tx$\rightarrow$Tc but also Tx and Tc are confirmed by E.

A similar situation applies to a refined method of testing of parameter-adjusted hypotheses that is widely used in statistics, namely *cross-validation*. Here one starts with just one (big) data set $E = (d_i{:}1\leq i\leq m\}$ (each $d_i$ being a data item), splits E randomly into two (disjoint) data sets $E_1$ and $E_2$, then fits the general hypothesis towards $E_1$ and tests the fitting result independently at hand of $E_2$. For each competing general hypothesis, one repeats this procedure several times and calculates its average likelihood ($\overline{P(E_2 | Tc_{E1})}$ of the second independent set. The result is highly reliable confirmation score. Another kind of method are the BIC (Bayes information criterion) and the AIC (Akaike information criterion); they are based on the probabilistic *expectation* value of the likelihood of a polynomial curve that is optimally fitted towards some set $E_1$, in regard to another independently chosen data set $E_2$ (cf. Hitchcock and Sober 2004).

*5.4 Uniting Worrall with Mayo and an Improvement of Mayo's Account.* According to Mayo's central idea, a test of a hypothesis H with outcome E confirms H (in a genuine way) iff the probability is low that *if H were false* the hypothesis would still pass this test with outcome E (Mayo 1991, 529; 1996, 274f). Mayo's intuition sounds exactly right, but the way she formulates it seems to be wrong. For as also Worrall (2010b, 151) pointed put, Mayo refers with "H" in her explication to the special theory Tc, i.e. the result of the accommodation of the general theory Tx to the evidence E. But both the data points E and the special theory Tc, i.e. the optimally E-fitted linear curve, are *fixed*; so even if Tc were false (which is improbable but possible by bad luck in data sampling) Tc would be in fit with E and hence would have passed the test with outcome E. What Mayo should refer to with the "hypotheses" in her criterion is not the special theory but the general theory Tx. In this reading, the test-procedure with outcome E includes both the adjustment of Tx's free parameters and the estimation of the fit of Tc with E (by the SSD-criterion). IF applied in this way, Mayo's criterion makes perfect sense: Tx (say the linearity assumption) would pass the test of being adjusted to E and then evaluated according to its fit with E, *even if Tx were false,* i.e., even if the true dependence between variables X and Y were not linear but say 3rd degree polynomial. Moreover, given this suggested improvement of Mayo's account, Worrall's objections to Mayo's account (2010b) lose their force. For, as we have argued in 5.3 above, if Tx can be fitted to every possible experimental outcome $E_i$, then $P(E_i|Tx) = P(E_i) = P(E_i|\neg Tx)$ must hold, which means exactly what the improved Mayo account says, namely that Tx would also pass the test, i.e. E would also have a high probability conditional on the assumption that Tx is false, i.e. ¬Tx is true. In conclusion, our improvements of Mayo's and Worrall's account are in perfect harmony.

## *6. Genuine Confirmation*

I turn now to my account of genuine confirmation, which consists in a uniform reconstruction of the achieved insights in terms of the probability increase of of *content elements*. The concept of confirmation is still the probabilistic – or if you want: the 'Bayesian' – one; what is new is that I apply this concept not only to the hypothesis H (or theory T) in toto but to the various *content parts*, or semantic meaning parts, of it. For the time being, think of content parts of a hypothesis H as of non-redundant consequences of H, and of *content elements* as 'smallest' content parts that are not further conjunctively decomposable – important refinements of this notion of content part will be given below 6.4. The idea of genuine confirmation is already found in Earman (1992, 106), but without explicaton. It rests on the observation that it is crucial for the notion of confirmation to be *evidence-transcending*: if we say that an evidence E confirms a hypothesis H that entails E then we mean that E does not only confirm H's content-part E but also those content-elements of H that go beyond E, i.e. are not entailed by E. We require this also in cases where H only raises E's probability but does not entail E. The requirement that *all* E-transcending content-elements of H should be confirmed by E is rather *strong*, and we consider weakenings of this requirement (to 'some' or 'some important' E-transcending content-elements) below.

We illustrate the concept of genuine confirmation at hand of successively more refined applications:

*6.1 Logical entailment:* If the hypothesis H is entailed by E ($E \Vdash\!\!- H$), then H doesn't have any E-transcending content-elements, whence the requirement that all E-transcending content elements of H are confirmed by H is trivially satisfied. So, there is no need to exclude entailment from confirmation: provided $P(H) < 1$, logical entailment is a limiting case of genuine confirmation (with $P(H|E) = 1 > P(H)$), in accordance with our intention.

*6.2 Tacking by conjunction* (cf. Glymour 1981, 67): Let E = grass is green and A be an absurd theory, e.g. the doctrine of Jehova's witnesses. Then mere conjunction of both, i.e. the hypothesis $H := E \wedge A$, is Bayes-confirmed by E (because it entails E, recall section 1, (2)). Surely, this probability-increase is not a case of genuine confirmation: here E is probabilistically irrelevant to the content-part of H that goes beyond E, namely A (i.e., $P(A|E) = P(A)$); E increases H's probability only because it is a content part of H and increases its own probability on trivial reasons to 1 ($P(E|E) = 1$). Gemes and Earman have called this type of non-genuine confirmation '*mere content-cutting*' (cf. Earman 1992, 98. 242, fn. 5; Schurz 2005, sec. 4).

*6.3 Inductive generalizations versus Goodman-generalizations*. The triviality of confirmation by mere content-cutting can also be seen from the fact that it applies to inductive as well as anti-inductive generalizations. For example, the evidence E that all so-far observed emeralds were green confirms via content-cutting the natural generalization $H_1$: "all emeralds are green" as well as the anti-inductive ('Goodman-type') hypothesis $H_2$ "all observed emeralds are green, and all non-observed ones are blue". Let $a_1,\ldots,a_n$ be the so-far observed emeralds; so $E = \forall x \in \{a_1,\ldots,a_n\}(Ex \rightarrow Gx)$. We can conjunctively decompose the hypothesis $H_1$ into the conjunction

(4) $H_1 = E \wedge H_1^*$, with $H_1^* = \forall x \notin \{a_1,\ldots,a_n\}(Ex \rightarrow Gx)$

and $H_2$ into the conjunction

(5) $H_2 = E \wedge H_2^*$. with $H_2^* = \forall x \notin \{a_1,\ldots,a_n\}(Ex \rightarrow Bx)$.

The content-part of $H_i$ that goes beyond E is in both cases $H_i$* (i = 1,2). Of course, only $H_1$* and not $H_2$* is inductively confirmed by E. Thus, only $H_1$ and not the Goodman-hypothesis $H_2$ is genuinely confirmed by E.[8]

That E confirms $H_1$* but not $H_2$* does, of course, *not* follow from the standard (Kolmogorovian) probability axioms alone (different from confirmation by content-cutting, which follows from them). In saying that E confirms $H_1$* but not $H_2$* I make the assumption that the probability measure P is *inductive*. Without this assumption content-transcending confirmation would be would be entirely impossible. A weak form of an inductive principle for P is the *exchangeability axiom* (explained in fn. 8); together with the regularity-axiom ("P(A) = 0 only if A is a contradiction") it entails that $P(Ga_{n+1} \mid Ga_1 \wedge \ldots \wedge Ga_n) > P(Ga_{n+1})$ and $P(\forall xGx \mid Ga_1 \wedge \ldots \wedge Ga_n) > P(\forall xGx)$ (cf. Earman 1992, 108).[9]

Of course, our confirmation intuitions go beyond inductive generalizations in the narrow (Humean) sense and include genuine confirmations of theories that contain *theoretical* concepts – i.e., concepts that are not contained in the evidence, and that were generated by processes of *abduction*. The question whether probability-increase of H by E spreads to H's E-transcending content-parts depends in this case on the use-novelty of E and is treated in below.

*6.4 The definition of content-elements*: The content-elements of a hypothesis H are all not arbitrary logical consequences. For example, "grass is green" is a content part

---

8 A further problem involved in "Goodman's new riddle of induction" (Goodman 1955) is *language-relativity:* by using the defined predicate "x is grue" $\leftrightarrow_{def.}$ "if x is observed, x is green, and otherweise blue", anti-inductive generalizations appear linguistically as inductive, and vice versa. I don't try to solve this problem in this paper but just assume a set of qualitative properties relative to which the notions of induction and anti-induction are defined.

9 Regularity implies $P(\forall xGx) > 0$. Without exchangeability, one cannot obtain induction results for finite evidence, but only inductive results *in the limit*, when the number of observed individuals approaches infinite. An induction axiom stronger than exchangeability is Carnap's indifference principle concerning structure descriptions (Carnap 1971).

of "apples are round and sweet", but "apples are round or made of cheese" is not a content part of "apples are round". The exclusion of irrelevant disjunctive weakening $H\lor X$ from being content parts of a hypothesis H is important in several areas of formal philosophy of science (cf. fn. 18). In our application, the admission or arbitrary irrelevant consequences as content parts would fall prey to the well-known "paradox" of Popper and Miller (1983) which runs as follows: every hypothesis H is logically equivalent with the conjunction of its consequences $(H\lor E)$ and $(H\lor\neg E)$, and while $(H\lor E)$ is already entailed by the evidence E, the second consequence $H\lor\neg E$ is provably always Bayes-disconfirmed by E, i.e. $P(H\lor\neg E \mid E) < P(H\lor\neg E)$. Miller concludes that non-deductive confirmation does not exist – a conclusion that is dubious on several reasons. But what his paradox clearly shows is that even if E Bayes-confirms H one may always found a logical consequence of H, namely $H\lor\neg E$, that is disconfirmed by E. But neither $H\lor E$ nor $H\lor\neg E$ are content-elements of E

In various papers I have developed a notion of content-element that has useful applications to many problems in formal philosophy[10] and provides a robust fundament for the concept of genuine confirmation. The definition goes as follows:

(6) *Def. 1: (1.1) S is a content element of H iff*

(a) H entails S,

(b) no predicate (including 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 $S_1\land\ldots\land S_n$ $(n\geq1)$ each of which is shorter than S.[11]

*(1.2)* S is a *content part* of H iff S is a non-redundant conjunction $S_1\land\ldots\land S_m$

[10] Cf. Schurz (1991, 1997, 2005), Schurz/Weingartner (2010). A congenial method of representation by content parts has been developed by Ken Gemes (1993, 2005). Gemes spoke of "content parts", but defined them differently than Schurz' "consequence elements" – a detailed comparison of both approaches is found in Schurz (2005, section 6).

[11] Length of a formula is defined as its number of primitive letters; and formulas are assumed to be transformed into their so-called *negation-normal form* (cf. Schurz/Weingartner 2010, sect. 4).

($m \geq 1$) of content-elements of H (i.e., no $S_i$ follows from the remainder conjuncts $\{S_j: j \neq i, 1 \leq j \leq m\}$).

Our notion of "part" admits "improper parts", i.e. T may be a content-part of itself. Here are some examples of so-called *irrelevant consequences* that violate condition (b) (the underlined occurrences are salva validitate replaceable): $p \Vdash p \vee \underline{q}$, $p \Vdash \underline{q} \rightarrow p$; $p \Vdash (p \vee \underline{q}) \wedge (p \vee \neg \underline{q})$; $\forall x(Fx \rightarrow Gx) \Vdash \forall x(Fx \rightarrow (Gx \vee \underline{H}x))$, $\forall x(Fx \rightarrow Gx) \Vdash \forall x((Fx \wedge Hx) \rightarrow Gx)$, etc. Examples of *relevant consequences* that do not violate condition (a), are: $p \wedge q \Vdash p$; $p \rightarrow q, q \rightarrow r \Vdash p \rightarrow r$; $Fa \Vdash \exists xFx$; $\forall x(Fx \rightarrow Gx), Fa \Vdash Ga$, $\forall x(Fx \rightarrow Gx), \forall x(Gx \rightarrow Hx) \Vdash \forall x(Fx \rightarrow Hx)$, etc.

Content elements (according to condition c) arise from relevant consequences by relevance-preserving conjunctive decomposition. Here are some *examples*, with $E(\Delta)$ for the set of content elements of a sentence set $\Delta$: $E(\{p \wedge \neg q\}) = \{p, \neg q\}$; $E(\{p \vee \neg q, p \vee q\}) = E(\{p\}) = \{p\}$; $E(\{p \rightarrow q, q \rightarrow \neg r\}) = \{\neg p \vee q, \neg q \vee \neg r, \neg p \vee \neg r\}$; $E(\{\forall x(Fx \rightarrow Gx), Fa\}) = \{\forall x(Fx \rightarrow Gx), Fa, \exists xFa, Ga, \exists xGx\} \cup \{Fa \rightarrow Ga: a \in Ic\}$ (where "Ic" = "set of individual constants of the language"), etc.

Two important facts about content elements are:

(A) *content-preservation:* For every sentence set $\Delta$, $E(\Delta)$ is always logically equivalent with $\Delta$, so no information gets lost by the representation of sentence sets (theories) by content elements (cf. lemma 7.2 of Schurz and Weingartner 2010); and

(B) *equivalence with clauses:* In propositional logic conten parts coincide with clauses (given the conventions in fn. 16; cf. Schurz/Weingartner 2010, lemma 5.1); even in predicate logical, an clause-theoretical reformulation is possible. This provides additional support, because clauses are an extremely well corroborated representation method in computer logic.

If an evidence E is a content part of H but H logically transcends E, then it need not always be the case (as in sect. 6.2 above) that H has a content part H* such that H is logically equivalent with $E \wedge H^*$ – but it follows from (A) above that in this case there must exist at least one content part of H that logically transcends E, and that must be Bayes-confirmed, too, if H is to be genuinely confirmed by E.

*6.5 Ex-post parameter adjustment:* According to the terminology introduced at the end of section 3, we have a background theory Tx = $\exists$xT(x) whose parameter(s) x is adjusted to a given evidence E, resulting in the special theory Tc (where c = $c_E$) that makes E highly probable. But since Tx could have been equally well fitted to every possible evidence $E_i$ (in the partition of possible evidences $\{E_i{:}i\in I\}$ that Tx intends to explain) in follows on ordinary probabilistic reasons that E cannot increase the probability of Tx (recall "improvement 3" of section 5). Since Tx is a content element of Tc that goes beyond E – indeed the most important content part of Tc that transcends E – the condition of genuine confirmation of $Tc_E$ by E is violated.

If Tx has been adjustéd by $E_1$ to Tc (c = $c_{E1}$) the crucial question is whether Tc makes some E-independent predictions $E_2$. This is not the case for rationalized creationism, but it is the case for curve fitting. Tx can of course not be fitted to any possible evidence $E_2$ via constants c that results from fitting of Tx to $E_2$-independent evidence $E_1$: by way of $E_1$-driven parameter-adjustment Tx makes only a small range of $E_1$-independent evidences probable, and if the observed $E_2$ lies within this range, then Tx's probability is raised by $E_2$, and Tc is genuinely confirmed by $E_2$.

Of course, the roles of $E_1$ and $E_2$ may be switched. So the confirmation of Tx now depends not only on the content of $E_i$ but also on the role it plays: whether it was used for parameter-adjustment or not. At this point the remark in section 2 becomes important, that the algebra of propositions over which the probability function is defined has to include also propositions describing procedures. Let $C_i$ be the procedural (or contextual) proposition that says that Tx was strengthened to Tc by adjusting x to $E_i$ (for i = 1,2). Then the precise conditional probability that we are confronted with are on the one hand $P(Tx \mid E_1 \wedge C_1)$ which equals P(Tx), and one hand $P(Tx \mid E_1 \wedge C_2)$, which is significantly raised above P(Tx), provided $E_1$ has confirmed Tc, i.e. $P(E_1|Tc)$ is significantly raised above $P(E_2)$.

*6.6 Non-genuine confirmation prevents sustainable probability-increase.* Even if Tc is not genuinely confirmed by E, because c was obtained from being adjusted towards E, the probabilistic fact remains that E increases Tc's probability (i.e., Bayes-confirms Tc). It is not an incoherence that E Bayes-confirms Tc abd Tx follows from Tc, although E doesn't Bayes-confirm Tx – it just means that E doesn't raise Tx's probability, i.e., $P(Tx|E) = P(Tx)$, where $P(T(x)) \geq P(Tc))$ because Tc $\Vdash$ Tx. But observe that *only one* content-part S of H is not confirmed by E and doesn't already have a high prior suffices *to prevent* sustainable probability-increase of Tc, in the sense that $P(Tc|E_1 \wedge E_2 \wedge \ldots)$ increases to successively higher values, when confirming evidences $E_i$ of the same type are accumulated. Since Tx's probability is not raised by any of these evidences $E_i$ (assuming that all of them are ex-post explained by parameter-adjustments), it holds that $P(Tx|E_1 \wedge E_2 \wedge \ldots) = P(Tx)$. But $P(Tx|E_1 \wedge E_2 \wedge \ldots)$ is an upper bound of $P(Tc|E_1 \wedge E_2 \wedge \ldots)$, because Tc $\Vdash$ Tx, whence $P(Tc|E_1 \wedge E_2 \wedge \ldots)$ is forced to stay below the low value of P(Tx), even if $i \rightarrow \infty$. This demonstrates the importance of the requirement that the Bayes-confirmation of a theory Tc by E can indeed spread to *all* parts of Tc. Only *then cumulative* confirmation and sustainable probability increase is possible, which is very important in science that works with theories whose prior probabilities are often very small. This insights justifies our characterization of genuine confirmation as *full* genuine confirmation, i.e. as Bayes-confirmation of all content parts of a H that go beyond E.

*6.7 Full versus partial genuine confirmation.* Despite the insight of section 6.6 it is desirable also to have a weaker notion of partial genuine confirmation. For it may often be that a theory T is (axiomatizable as) a conjunction $T_1 \wedge T_2$ such that $T_1$ but not $T_2$ is confirmed by E – for example, because $T_1$ (together with given background conditions) entails E, while $T_2$ is not needed in this entailment. There are many historical examples of successful theories containing empirically useless assumptions, such as the assumption that the sun is the centre of the universe in Newtonian celestial mechanics. One may even add to an established scientific theory $T_1$ a creationist postulate $T_2$ (which is sometimes done by rationalized creationist; recall sect. 1). In these cases we want to say that E partially genuinely confirms T because at least some content part of T, namely $T_1$, is genuinely confirmed. It is important to insist, however, that T's content part $T_1$ must be fully genuinely confirmed by E, i.e. the probability of each content element of the content part $T_1$ must be increased – because only then, by the insight of section 6.6, can $T_1$'s probability be sustainably increased by accumulating evidence. Moreover, if Bayes-confirmation of some content part were enough for partial genuine confirmation, the latter notion would collapse into Bayes-confirmation, because by fact (A) of sect. 6.5, every theory is logically equivalent with some content part of itself. For example, the ex-post theory T = "God created E" has as its content elements {"God created E", E, $\exists$x(x created E), $\exists$X(God created X), $\exists$x$\exists$X(x created X)}, and as content parts all non-redundant conjunctions of these elements. The last three (existentially quantified) content elements are contain in the first, "God created E", but the are not Bayes-confirmed by E, so neither "God created E" nor any other E-transcending content part of T is fully genuinely confirmed by E, whence T is not even partially confirmed by E.

If the hypothesis H is not a mere inductive generalization but a proper theory that contains *theoretical* concepts or parameters[12], then the given characterization of partial genuine confirmation has to be strengthened to the effect that the confirmation of the theory should spread at least to *some* of the theoretical parameters. For otherwise the theoretical structure is entirely superfluous and should be abandoned, in line with Ockham's razor. In this case one should not say that the theory is confirmed, not even 'partially' – what is only confirmed in this case is the theory's empirical content.

We illustrate this at hand of two examples. As a first example assume an intelligent design theorist explains a particular evidence E by a creationist hypothesis H that does not only entail E but an inductive generalization of E that generates novel predictions E':

[12] It is hotly debated whether the distinction between theoretical and non-theoretical concepts is historically relative or non-relative, but we need not take a stance on this question.

| | | |
|---|---|---|
| (7) | Observed evidence E: | So far the sun was rising every day |
| | is explained by T: | God makes it that the sun rises every day |
| | entailing 'novel' prediction E': | The sun will rise also in the future, |

In spite of the fact that E confirms now E', which is an E-transcending content part of T, we should not consider this as a case of genuine confirmation of the *theory* T, because that the sun raises every day (G) is a mere inductive generalization of E that is sufficient to predict the novel prediction E'. The postulate of the theoretical entity "God" is superfluous for this inductive prediction. In terms of ex-post adjustment, the general theory Tx = "there exists so-far observed regularities $\psi(t)$ that will continue in all future because God takes care for their continuation" may be adjusted to every possible so-far observed regularity E(t) (by replacing $\psi(t)$ by E(t)), whence its probability is not increased by E(t). So we may say, theory T in (7) is not genuinely confirmed by E' because E' does not increase the probability of any E-transcending content-part of T that entails the existence of the theoretical entity "God".

Compare example (7) with the following example of a scientifically justified theoretical explanation:

| | | |
|---|---|---|
| (8) | Observed evidence E: | So far the sun was rising every day |
| | is explained by T: | The earth rotates with one rotation per day |
| | entailing 'novel' prediction E': | All stars turn over the nightly horizon each day with equal rotating speed. |

In (8) the prediction E' *qualitatively* novel: E' is not inferable from E by mere induction. The theoretical assumption of the earth's rotation is essential for deriving E', from which we may conclude that the probability of the theoretical content parts of T is increased by E', i.e., T is genuinely confirmed by E'.

Our second example is again curve fitting of data E by a linear curve Tc via an optimal adjustment of the general linearity hypothesis Tx. Recall from sect. 3, what is Bayes-confirmed by E in this case is neither Tc nor Tx but the implication Tx$\rightarrow$Tc. Is therefore Tc at least partially genuinely confirmed by E? First of all, Tx$\rightarrow$Tc is an irrelevant consequence of Tc and thus not a content part (Tx is salva validitate re-placeable by arbitrary predicate $\psi$x). However there exist a logically equivalent re-formulation of Tx$\rightarrow$Tc that is a relevant consequences, even a content-element, namely (applying the terminology of sect. 4):

(9) $\forall\alpha_o\forall\alpha_1\forall\sigma$ [ $(\forall d(Y(d) = \alpha_1\cdot X(d)+\alpha_0+r(\sigma)) \rightarrow \alpha_1 = c_1 \wedge \alpha_0 = c_0 \wedge \sigma = s$ ]

*In words*: For all $x_i$, *if* Y is linearly dependent on X with parameter values $x_i$, then the parameter-values $x_i$ are identical with constants $c_i$.

The content element (9) of Tc transcends E and is moreover fully genuinely confirmed, because all content element it contains are Bayes-confirmed by E. However, $\sigma_1, \alpha_0$, $\sigma$ are *unobservable* and hence *theoretical* parameters that define a linear (theoretical) function, and the content element (9) does *not* entail the existence of a linear X-Y-dependence, i.e., the existence claim $\exists\alpha_o\exists\alpha_1\exists\sigma$ [$\forall d(Y(d) = \alpha_1\cdot X(d)+\alpha_0+r(\sigma)$] is not derivable from (9). Moreover no other fully genuinely confirmed content part of Tc exists that would entail the existence of a linear X-Y-dependence. Therefore we conclude that Tc is *not* partially genuinely confirmed by E, in accordance with out intuitions. We summarize our final explication of genuine confirmation in the following definition:

(10) *Def. 2: (2.1)* A hypothesis H is (fully) genuinely confirmed by an evidence E iff every content element S of H that is not logically contained in E is Bayes-confirmed by E ($P(S|E) > P(S)$).

*(2.2)* A hypothesis H is partially genuinely confirmed by an evidence E iff some content part H* of H that is not logically contained in E is Bayes-confirmed by E

(P(S|E) > P(S)), and moreover, if H contains theoretical concepts (including parameterized functions), then H* entails the existence of at least one theoretical concept φ of H.[13]

*6.8 Criteria for the spread of probability-increase to a theory's content parts.* So far we have given several reasons that prevent the spread of probability increase of T by E to certain content parts of T. But what positive reasons do we have for assuming a spread of probability-increase? An orthodox Bayesian would answer, this is determined by the relation between the prior and posterior probabilities of the content parts of T; however typically the priors of scientific hypothesis are completely undetermined, and even more the priors of the content parts. For example, what is the prior probability of the existence of anti-gravity, life on Mars, of the development of eukaryotes from prokaryotes? – no scientist that I know asks such questions because the answer could only be arbitrary.

The question whether the probability of a content part S of T is raised by E, given T's probability is raised by E, should rather be answered according to the role and *weight* this content part S played in the increase of the probability of E by T – this weight should be an indicator of probability increase of E by S, and thus of S by E. The weight of content part S is partly dependent on the concrete nature of T and of the given relevant background knowledge – hence there exist no sufficient general answer to this question. But there exist the following necessary general criteria for spread of probability-increase:

(11) *Necessary criteria for spread of probability-increase:* If T increases E's probability, then the resulting probability increase of T by E (P(T|E) > P(T) spreads from T to an E-transcending content part S of T only if:

---

[13] If c is a parameterized function $f(c_i)$, the existence claim is 1st order and has the form $\exists\alpha_i(\alpha_i{=}c_i \wedge A[f(\alpha_i)])$; if c is a theoretical property F, the existence claim is 2nd order and has the form $\exists\varphi(\varphi = F \wedge A[F])$.

(1) S is necessary within T to make E highly probably, i.e., there exists no content part T* of T that does not entail S but makes E equally probable than T (P(E|T*) = P(E|T)), and
(2) T does not result from a parameter-adjustment of S to E that would have been equally possible for every possible experimental outcome $E_i$.

(2) plays the role of a *defeater* that blocks spread of probability increase; further context-dependent defeaters may be added to this explication (e.g., that E's data intems have not been selected in favor of T).

*6.9 Conclusion.* In conclusion, the developed concept of (fully or partial) genuine confirmation has the following advantages:

(a) It provides answers to the objections that have been raised against Worrall's account of use-novelty, by offering three improvements of this account without deviating from its spirit;

(b) it offers an improvement of Mayo's account of severe tests that is in harmony with the use novelty criterion,

(c) it is uniformly applicable to the confirmation of inductive generalizations, curve fitting, and proper theories (including entailment as the extreme case of confirmation), and finally

(d) it develops purely probabilistic justifications of prima facie "non-Bayesian" confirmation criteria by considering the probability-increase of evidence-transcending content parts of hypotheses. This leads to a concept of genuine confirmation that unifies Worrall's use-novelty account and Mayo's severe test account with a refined probabilistic account to confirmation.

*Acknowledgement:* Work on this paper was supported by the DFG. For valuable discussions on the topic of this paper I am indebted to John Worrall, Deborah Mayo, Ludwig Fahrbach and Franz Huber.

*References*

Behe, M. (2003): "The Modern Intelligent Design Hypothesis: Breaking Rules", in: Manson, N.A. (2003, eds.), 276-290.

Carnap, R. (1971), "Inductive Logic and Rational Decisions", in: Rudolf Carnap and Richard Jeffrey (Eds.), *Studies in Inductive Logic and Probability I*. Los Angeles : Univ. of Calif. Press.

Dembski, W. (2003): "The Chance of the Gaps", in: Manson, N.A. (2003, eds), 251-327.

Earman, J. (1992), *Bayes or Bust?* Cambridge/Mass.: MIT Press.

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

Gemes, K. (2005). "Hypothetico-Deductivism: Incomplete but not Hopeless", *Erkenntnis* 63, 139-147.

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

Goodman, N. (1955): *Fact, Fiction and Forecast,* 2005 Harvard Univ. Press, Cambridge/MA

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.).

Isaak, M. (2002): „What Is Creationism?“, http://www.talkorigins.org/faqs/wic.html. (links updated: Dec 12, 2002).

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.

Laudan, L. (1983): "The Demise of the Demarcation problem", in: R.S. Cohen, and L. Laudan (eds.), *Physics, Philosophy, and Psychoanalysis*, Reidel, Dordrecht, 111-127.,

Manson, N.A. (2003, eds), *God and Design*, Routledge, London and New York.

Mayo, D. (1991): "Novel Evidence and Severe tests", *Philosophy of Science* 58, 523-552.

Mayo, D., and Spanos, A. (eds., 2010): *Error and Inference*, Cambridge Univ. Press, New York.

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

Pennock, R.T. (2011). "Can't Philosophers Tell the Difference between Science and

Religion? Demarcation Revisited", *Synthese* 178, 177-206.

Popper, K., and Miller, D. (1983): "A Poof of the Impossibility of Inductive Probability", *Nature* 302, 687-688.

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

Schurz, G. (1997): *The Is-Ought Problem*, Kluwer, Dordrecht.

Schurz, G. (2005): "Kuipers' Account to H-D Confirmation and Truthlikeness", in: Festa, R. (Hg.), *Logics of Scientific Discovery. Essays in Debate With Theo Kuipers*, Rodopi, Amsterdam.

Schurz, G. (2008): "Patterns of Abduktion", *Synthese* 164, 201-234.

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.

Watkins, J.W.N. (1964): "Confirmation, the Paradoxes, and Positivism", in: M. Bunge (ed.), *The Critical Approach to Science and Philosophy*, xx, 92-115.

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

Worrall, J. (2010b): "Error, Tests, and Theory Confirmation", in: D. Mayo and A. Spanos (eds., 2006), 125-154.

Zahar, E.G. (1973): "Why did Einstein's Programme supersede Lorentz's", *British Journal for the Philosophy of Science*, 24, 95-123, 223-262.