From the indirect confirmation of theories to theory unification

ABSTRACT. Theory unification is a central aim of scientific investigation. In this paper, I lay down the sketch of a Bayesian analysis of the virtue of unification that entails that the unification of a theory has direct implications for the confirmation of the theory's logical consequences and for its prior probability. This shows that scientists do have epistemic, and not just pragmatic, reasons to prefer unified theories to non-unified ones.

1. Introduction

Theory unification is a central aim of scientific investigation; topical examples are the recurring attempts to provide, in physics, a unified theory of both quantum and relativistic phenomena. It is therefore important for the methodology of science to give an account of the property of unification. One possible view is that scientists prefer unified theories to non-unified rivals for just *pragmatic* reasons (for instance, because unified theories are simpler and thus easier to 'handle'). Indeed, this would seem to follow from van Fraassen's conception of a theory's explanatory power and non-empirical epistemic virtues in general.[1] Against views of this kind, I will argue that scientists do have *epistemic* reasons to prefer unified to non-unified theories, as they believe that unification increases their reliance on the scientific statements that they accept. To show it, I will lay down the sketch of a Bayesian analysis of unification that entails that the unification of a theory has direct implications for the confirmation of the theory's logical consequences and for its prior probability.

Intuitively, a theory is unified if it is not a mere collection of disparate hypotheses but a system of laws and principles that connect to one another. For example, it is intuitive that such principles should be as few as possible (the minimum necessary to explain all relevant evidence). They should be logically consistent with one another and similar to each other in terms of logical structure. Finally, they should share non-logical predicates. (See, for instance, Watkins, 1984, pp. 203-215 and Causey 1977, pp. 111-121). No matter how unification is explained in details from a logical point of view (for the purpose of the paper I assume that such a detailed explanation can be provided), there are properties that scientists ascribe to unified theories when they look at them from an epistemic perspective. My purpose is not to try to justify the *correctness* of those ascriptions (for instance, from a realist point of view) but, rather, to clarify them by means of a mathematical formulation of their content.

---

[1] See Van Fraassen 1980.

In the next two sections, I will outline an account of unification resting on the explanation of the fact that unified theories are used, in science, for the indirect confirmation of their theoretical and observational consequences, and resulting in a necessary Bayesian condition that a unified theory must meet. In the conclusion, I will present a possible case of theory choice based on the probabilistic consequences of unification.

2. Indirect confirmation and unification

There are direct links between unification and empirical confirmation. These links result, to begin with, in the fact that unified theories allow their *theoretical* consequence to be indirectly confirmed by observations.

The notion of indirect confirmation can be spelled out as follows: let *T* be a scientific theory that entails a hypotheses *H*. Let us suppose that *T* also entails an observational statement *E*, which is not logically derivable from *H* (i.e. neither from *H* alone nor from *H* in conjunction with a set of auxiliaries).[2] The advocates of indirect confirmation maintain that, in such a situation, if *E* confirms *T*, the confirmation can be 'communicated' to *H* through the 'bridge' theory *T*. Thus, *E* indirectly confirms *H*. (See, for instance, Laudan and Leplin 1991). Indirect confirmation has effective uses in science. In physics, a celebrated example is given by the indirect support for the kinematic laws (*H*) of the Theory of Relativity (*T*) obtained by the achieved increase of mass with velocity (*E*). (For this and other examples, see Laudan and Leplin 1991 and Laudan 1995).

As Okasha (1997) has argued, the indiscriminate use of the method of indirect confirmation yields paradoxical conclusions. Suppose in fact that the method is applicable in *any* case. Namely, to the effect that, given an observational statement *E* and a non-tautological hypothesis *H* not entailing *E*, *E* indirectly confirms *H* if there is a logically consistent theory *T* that entails both *E* and *H* and is confirmed by *E*. Considering that (on basic hypothetico-deductivist principles) *T* is confirmed by a logical consequence *E* if and only if *T* is logically consistent and *E* is not a tautology,[3] the above condition can be re-written as follows:

(IC1) If *E* is a non-tautological observational statement and *H* is a non-tautological theoretical statement that does not entail (alone and in conjunction with auxiliaries) *E*, *E* indirectly confirms *H* if there is a logically consistent theory *T* that entails *E* and *H*.

[2] Notice that *T* cannot be considered an auxiliary hypothesis of *H*, as *T entails H*. Moreover, if *E* is derivable from another theoretical consequence *H** of *T*, *H** cannot be considered an auxiliary hypothesis that enables *H* to entail *E*. For *H**, in this case, entails *E* also when *isolated* from *H*.

[3] I include under the term 'tautology' also conventional definitions and analytic statements.

It is easy to show that (IC1) is untenable. Let *E* be the statement that, in observed cases, mass does increase with speed, and let *H* be the hypothesis of the existence of unconscious process in the human minds. Finally, let *T* be the mere *logical conjunction E* & *H*. In this case, *E* and *H* are not tautologies and *T* is logically consistent. Since *T* entails both *E* and *H*, and *H* does not entail *E*, on the grounds of (IC1), *E* indirectly confirms *H*. But nobody – and no scientist – would ever accept it, as *T*'s confirmation by *E* appears, in this case, *irrelevant* for the confirmation of *H*.

How should we explain this irrelevance? In my opinion, an insightful answer is the following. It is only by the confirmation of the general principles that unify a set of statements in one theory that evidential support is 'transmitted' to the hypotheses that do not entail the confirmatory evidence. Since *T* includes no general principles able to unify *E* and *H*, and *H* does not entail *E, T*'s confirmation by *E* cannot 'flow' to *H*. (This view is defended in Moretti 2002). This answer appears to match and to provide explanation of the actual scientific practice, where unified theories are typically used to 'channel' indirect confirmation.[4]

In conclusion, a rule for the application of the indirect confirmation method that apparently reflects the actual scientific practice is following:

(IC2) If *E* is a non-tautological observational statement and *H* is a non-tautological theoretical statement that does not entail (alone and in conjunction with auxiliaries) *E*, *E* indirectly confirms *H* if there is a *unified* theory *T* that entails *E* and *H*.

(Notice that, since unification certainly entails logical consistency, the specification that *T* must be logically consistent is dispensable). (IC2) enables scientists to circumscribe the applicability range of the indirect confirmation method so that to rule out the counterintuitive confirmation cases just considered.

In science and in everyday life, there are cases of indirect confirmation that apply to simple empirical generalisations. They depend on the trivial inductivist belief according to which, if a statement $Pa \supset Qa$ confirms a *nomic* generalisation $\forall x(Px \supset Q\mathrm{x})$, then confirmation 'spills' over the parts of $\forall x(Px \supset Q\mathrm{x})$ that extend beyond $Pa \supset Qa$ – for instance, $Pb \supset Qb$. (Cf. Gemes 1998). In such cases, the relation of indirect confirmation holds not between an observational and a

[4] Typical examples of it are the indirect support for Kepler's laws of planetary motion via the independent confirmations of Newton's mechanics (see Laudan 1995) and the indirect support for the Special Relativity Theory via the independent confirmations of the General Relativity Theory (see Newton-Smith 1980). It is worth emphasising that Newton's Mechanics and the General relativity Theory are just *paradigmatic* examples of theory unification.

theoretical sentence but between two observational sentences – for instance, $Pa \supset Qa$ and $Pb \supset Qb$. In these cases too, the appeal to the unifying power of nomic generalisations allows the inductivist to get rid of superabundant and counterintuitive cases of confirmation.

This consideration makes it plausible that two observational statements confirm each other if they are deducible from a unified theory (which can consist in a mere set of nomic generalisations). We can then generalise (IC2) as follows:

(IC3) If *E* is a non-tautological observational statement and *C* is a non-tautological statement that does not entail (alone and in conjunction with auxiliaries) *E*, *E* indirectly confirms *C* if there is a unified theory *T* that entails *E* and *C*.

This is the condition that apparently governs the main applications of the indirect confirmation method in the actual scientific practice.

3. A Bayesian analysis of unification

How should we explain, in general, the fact that scientists assign to unified theories the methodological function described by the rule (IC3)? I put forward the following explanation that strikes me as very intuitive: the belief that grounds rules like (IC3) is – roughly – that *any* part of a unified theory is *epistemically* tied to *any* of the other parts in virtue of the theory's unifying principles. This means, more precisely, that every non-tautological consequence of a unified theory confirms every other non-tautological consequence of the theory and that – equivalently – every non-tautological consequence of a unified theory is confirmed as the theory is confirmed. This is the ultimate reason why – on my view – theory unification provides the very grounds of the 'mechanism' of indirect confirmation used in science.

What I said singles out the following necessary *epistemic* condition for theory unification:

(U1) If *T* is a unified theory, for every non-tautological consequences *C* and *C** of *T*, *C* confirms *C**.

Notice that (U1) is trivially equivalent to:

(U2) If *T* is a unified theory, for every consequence *C* of *T* and any non-tautological consequence *C** of *T*, if *C* confirms *T*, *C* confirms *C**.

Notice finally that (U1) and (U2) entail (IC3).

Let us now 'translate' (U1) and (U2) into quantitative conditions. Let A and B be two statements and let K be background knowledge. In accordance with the Bayesian definition of incremental confirmation, A confirms B relative to K if and only if Pr $(B|A \,\&\, K) >$ Pr $(B|K)$. In what follows, for the sake of simplicity, I will assume that, for every statement A, Pr $(A|K) = 1$ if and only if A is a tautology and Pr $(A|K) = 0$ if and only if A is logically inconsistent.

The Bayesian definition of confirmation and the above assumption allow us to 'translate' (U1) into the following probabilistic condition:

(BU1) If T is a unified theory,[5] for every logical consequence C and C^* of $T \,\&\, K$ such that Pr $(C|K) < 1$ and Pr $(C^*|K) < 1$, Pr $(C^*|C \,\&\, K) >$ Pr $(C^*|K)$.

(BU1) expresses the insight that, if a theory is unified, its non-tautological consequences are probabilistically dependent on one another in a positive sense. This claim is the quantitative counterpart of the qualitative one, made by (U1), according to which, if a theory is unified, its non-tautological consequences confirm each other.

By analogy with the relation existing between (U1) and (U2), (BU1) is logically equivalent to:

(BU2) If T is a unified theory, for every logical consequence C and C^* of $T \,\&\, K$ such that Pr $(C^*|K) < 1$, if Pr $(T|C \,\&\, K) >$ Pr $(T|K)$, then Pr $(C^*|C \,\&\, K) >$ Pr $(C^*|K)$.

(BU2) expresses in quantitative terms the claim, made by (U2), that if a theory is unified, whenever the theory is confirmed, any of its non-tautological consequences is confirmed too.[6]

[5] Thus Pr $(T|K) > 0$.

[6] Notice that it is not completely implausible that (BU1) and (BU2), as they stand, might turn out to be necessarily false. The reason is that, for any logical conjunction $T \,\&\, K$ such that Pr $(T|K) > 0$, if the non-empty set **S** of all the logical consequences C of $T \,\&\, K$ such that Pr $(C|K) < 1$ does exist, **S** might include couples of statements (C, C^*) for which it is *probabilistically incoherent* that Pr $(C^*|C \,\&\, K) >$ Pr $(C^*|K)$. I cannot address this complicated problem in so short a paper. I wish however to emphasise that, if this possibility proved actual, it could possibly be settled by a refinement of (BU1) and (BU2) – and thus of (U1) and (U2) – that specifies an opportune sub-set **S'** of **S** such that for any couple (C, C^*) of the members of **S'** it must be true that Pr $(C^*|C \,\&\, K) >$ Pr $(C^*|K)$ when T is unified. It seems to me that any such restriction would still harmonize with the spirit of the present account of unification if it would not prevent any couple (C, C^*) of members of **S** that are intuitively *part of the content* of $T \,\&\, K$ from satisfying the condition that Pr $(C^*|C \,\&\, K) >$ Pr $(C^*|K)$. (For a formal definition of theory content, see for instance Gemes 1998).

4. Conclusion

The notion of theory unification plays a tangible role in scientific methodology, as unified theories are part of the framework by which theoretical and observational statements can receive indirect support by evidence. This suggests a reason why scientists typically prefer unified theories to non-unified rivals: the former, but not necessarily the latter, provide scientists with an additional device to increase a scientist's reliance on certain scientific statements that he believes in (for instance, a scientist's confidence in a hypothesis H will increase as soon as H is embodied in a confirmed unified theory). This shows that the preference of scientists for unified theories depends, very plausibly, on also *epistemic* not on just pragmatic reasons.

Indeed, there is another way to argue for the same conclusion that hinges on the analysis of unification given in terms of the conditions (BU1) and (BU2). I will now show that, on certain circumstances, the application of (BU1) enables a unified theory to raise its prior probability over the prior probability of its non-unified rivals, and this makes the theory epistemically preferable to its rivals. Hence, insofar (BU1) reflects with sufficient accuracy the actual conception of unification of scientists, this principle shows that, at least in certain cases, scientists will prefer unified theories on non-unified rivals on genuine epistemic grounds.

Let suppose that a unified theory T, available at a time t, is developed, at a time t+n, into two incompatible rivals T_1 and T_2. T_1 is still unified, as it consists of the conjunction T & H_1, where H_1 is a unified set of hypotheses homogeneous with T. Quite the opposite, T_2 is no longer unified. For it consists of the conjunction of T & H_2, where H_2 is a set of hypotheses including laws and principles very dissimilar (for instance, in terms of their logical structure) from those included in T. Let us moreover suppose that H_1 and H_2 are equally plausible given background knowledge K, so that one can stipulate that (a) $\Pr(H_1|K) = \Pr(H_2|K)$. Finally, let us assume that (b) $\Pr(T_1|K), \Pr(T_2|K) > 0$ and $\Pr(T|K), \Pr(H_1|K), \Pr(H_2|K) < 1$. I will now show that, given these conditions, if (BU1) is accepted and a very intuitive assumption is made (see below), then $\Pr(T_1|K) > \Pr(T_2|K)$.

Notice that all conditions to apply (BU1) to T_1, T and H_1 are fulfilled, as T_1 is unified, T and H_1 are both entailed by T_1 & K and, because of (b), $\Pr(T|K) < 1$ and $\Pr(H_1|K) < 1$. Since (BU1) applies, it follows that $\Pr(T|H_1 \& K) > \Pr(T|K)$, which trivially entails that (1) $\Pr(T \& H_1|K) > \Pr(T|K)\Pr(H_1|K)$.

In contrast, since T_2 is not unified, as its parts T and H_2 are not homogeneous, there is no reason to believe that $\Pr(T|H_2 \& K) > \Pr(T|K)$, while it is quite intuitive that $\Pr(T|H_2 \& K) \leq \Pr(T|K)$. This entails that (2) $\Pr(T \& H_2|K) \leq \Pr(T|K)\Pr(H_2|K)$.

From (1), (2) and the assumption (a) that $\Pr(H_1|K) = \Pr(H_2|K)$, it follows that (3) Pr (T &

$H_1|K) > \Pr(T \,\&\, H_2|K)$. Since $T \,\&\, H_1$ and $T \,\&\, H_2$ are respectively equivalent to T_1 and T_2, (3) is equivalent to $\Pr(T_1|K) > \Pr(T_2|K)$. QED.[7]

Luca Moretti

Philosophy, Probability, and Modeling Research Group

Center for Junior Research Fellows

University of Konstanz

Luca.Moretti@uni-konstanz.de

References

Causey, R. *Unity of Science*. Dordrecht: Reidel.

Gemes, K. 1998. Hypothetico-Deductivism: the Current State of Play; the Criterion of Empirical Significance: Endgame. *Erkenntnis* 49: 1-20.

Laudan, L. 1995. Damn the Consequences! *Proceedings and Addresses of the American Philosophical Association* 69: 27-34.

Laudan, L. and Leplin, J. 1991. Empirical Equivalence and Underdetermination. *The Journal of Philosophy* 88: 449-72.

Moretti, L. 2002. For a Bayesian Account of Indirect Confirmation. *Dialectica* 52: 153-73.

Newton-Smith, W. 1980. The Underdetermination of Theory by Data. In Hilpinen, R. (ed.), *Rationality in Science*. Reidel: Dordrecht.

Okasha, S. 1997. Laudan and Leplin on Empirical Equivalence. *The British Journal for the Philosophy of Science* 48: 251-56.

Van Fraassen, B. 1980. *The Scientific Image*. Oxford: Clarendon Press.

Watkins, J. W. N. 1984. *Science and Scepticism*. Princeton: Princeton University Press.

[7] I wish to thank Phillip Meadows for comments and suggestions upon early versions of this paper.