Occam's Razor in the Theory of Theory Assessment

August Fenk, Klagenfurt, Austria

1. Overview

In this paper I will at first discuss the role of economy, parsimony or simplicity in theory assessment and model selection. This discussion (in Section 2) will amount to a three-dimensional model of theory assessment, including Coombs' (1984) dimensions generality (breadth) and power (depth), and simplicity as the third dimension.

Theory assessment is, most commonly, a matter of the methodology of empirical science. But its principles might also apply to "metaphysical theories", at least in part, as already suggested in Laszlo (1972:389). Thus they might also be applicable, in selfreferential ways, to those meta-theory – the "theory of theory assessment" in terms of Huber (2008:90) – that has invented the above mentioned criteria of model selection and theory assessment. This is exactly what I shall study in Section 3 of this paper, focusing on the key-concepts of *law* and *lawlikeness.* Laws are usually assumed to be a precondition for the reconstruction and explanation of phenomena on the one hand and their anticipation and prediction on the other, but relative frequency will be shown as the proper basis of all our projections to the past and to the future. Evolutionary perspectives are indicated in the last Section 4.

Thus, this paper does not deal with the reduction of theories in the sense of Nagel (1961), or with the problems in the attempts to reduce "emergent" systems to their elements, but rather with the *reduction* of (semantic) complexity and the *elimination* of dispensable components of (meta-)theories. And, in a certain sense, with the "reduction" of *law* to statistical generalizations.

2. Three dimensions of theory assessment

Most theories of theory assessment are two-dimensional, balancing e.g. "empirical adequacy" against "integrative generality" (Laszlo 1972:388) or power against generality (Coombs 1984), and most of the standard methods of model selection provide, according to Forster (2000:205), "an implementation of Occam's razor, in which parsimony or simplicity is balanced against goodness-of-fit".

But there are also some attempts to three-dimensional models: In his above mentioned paper Forster (2000:205) suggests that model selection should, besides simplicity and fit, "include the ability of a model to generalize to predictions in a different domain". In Lewis (1994:480) there is talk about a trade off between the "virtues of simplicity, strength, and fit". And Laszlo's (1972:388) factor "integrative generality" figures as "a measure of the internal consistency, elegance, and 'neatness' of the explanatory framework". Two scientific theories, he says, can be compared with regard to the number of facts taken into account (I), the precision of the accounting (II), and the economy (III) whereby the balance between "integrative generality" and "empirical adequacy" is produced. Economy (III) is, first of all, associated with a small number of "basic existential assumptions and hypotheses" (Laszlo 1972:388). (I) and (II) correspond to Coombs' generality and power, and Coombs' model may be viewed as an appropriate decomposition of Laszlo's factor "empirical adequacy". But it fails to account for Occam's razor.

Considering such arguments I emphasize a three-dimensional model (Fenk 2000) including the dimensions precision, generality (size of domain), and parsimony, as well as a strict distinction between the theory's assertions – the lawlike propositions in the core of any scientific theory – and the theory's "predictive success" (in the sense of Feyerabend 1962:94). Other than in the above mentioned approaches by Forster and by Lewis, goodness-of-fit is not a separate dimension, but the touchstone of the whole theory. According to this model we state an advantage of a theory t2, as compared with a former version or conflicting theory t1, if it achieves at least the same predictive success (number of hits) despite a higher precision of the predictions and/or an extended domain and/or a lower number of assumptions. With regard to Coombs' trade-off between the dimensions „power" and „generality", this idea is illustrated in Fenk & Vanoucek (1992:22f.), though only on the level of single lawlike assumptions.

Popper (1976:98,105) suggests disregarding, at least in epistemological contexts, properties of pure representation as well as the respective conventionalistic, "aesthetic-pragmatical" conceptualizations of "simplicity" or "elegance". But maybe the aesthetic attributes come by the theory's economic functionality, just as in the aesthetic BAUHAUS-principle "form follows function"? And our three-dimensional model actually applies, first of all, to *theory* as a hypothetical representation or construction. It is particularly interesting to see that it none the less fits all of Popper's further arguments regarding the relations between "empirical content", "testability", and "simplicity": The more possibilities ruled out by a sentence ("je mehr er verbietet"; p. 83), the higher its empirical content. "Auf die Forderung nach möglichst großem empirischen Gehalt können noch andere methodologische Forderungen zurückgeführt werden; vor allem die nach möglichst großer *Allgemeinheit* der empirisch-wissenschaftlichen Theorien und die nach größter Präzision oder *Bestimmtheit*." (p. 85) „Einfachere Sätze sind /…/ deshalb höher zu werten als weniger einfache, weil sie *mehr sagen*, weil ihr empirischer Gehalt größer ist, weil sie besser überprüfbar sind." (p. 103) Thus, generality (Allgemeinheit), precision (Bestimmtheit) and simplicity (Einfachheit) turn out to be three different facets of Popper's essential idea of testability and the chance to be falsified.

Are virtues such as "integrative generality" and "economy", as suggested in Laszlo (1972:389), also applicable to "metaphysical" disciplines, i.e. to meta-theories that have to do without the corrective of direct empirical tests? In theoretical semiotics, for instance, a reduced complexity of the terminological framework may allow to solve classificational problems such as the definition of *iconicity* (Fenk 1997), or to solve and communicate them in better understandable ways.[1] Can we apply criteria of scientific progress invented by the philosophy of science even to essential concepts of that philosophy of science?

---

1 The latter aspect reminds, in some ways, of the concepts of "userfriendlyness" in Cognitive Ergonomics and of (low) "item-difficulty" in test theory.

3. A reductionistic look on laws and lawlikeness

A general principle „that is applicable to all kinds of reasoning under uncertainty, including inductive inference" (Grünwald 2000:133) – is such a thing conceivable in view of the problems discussed in the philosophy of science?

I will try that focusing on the key-concepts of *law* and *lawlikeness*. In Goodman (1973:90,108) a hypothesis is lawlike only if it is projectible and projectible when and only when it is supported (some positive cases), unviolated (no negative cases), and unexhausted (some undetermined cases)[2]. But especially the criterion "unviolated" seems to be rather meant for universal laws (Fenk & Vanoucek 1992). What should be considered the negative and the positive cases in view of a weak regularity such as a very severe side-effect of a medicament showing in one of hundred patients in nine of ten studies?

The following outline starts with the universal laws in the Deductive-Nomological (D-N) model by Hempel & Oppenheim (1948). The authors note that their formal analysis of scientific explanation applies to scientific prediction as well. This symmetry between explanation and prediction will outlast. The application of the D-N model, however, is restricted to a world of universal laws – a rather restricted or even non-existent world, if *law* is not understood as a mere proposition but as an empirically valid argument. Thus we see a shift of the focus in the philosophy of science from the universal laws in the D-N model to statistical arguments rendering their extremely high probabilities ("close to 1") to the explanation in Hempel's (1962) Inductive-Statistical (I-S) model. And from here to the reduction of "plausibility" to the relative frequencies observed so far (Mises 1972:114) and to "stable" frequency distributions as a sufficient basis for "objective chances" (Hoefer 2007). Let me carry that to the extremes: If a dice had produced an uneven number in ten of fifteen cases I would, if I had to bet, bet on "uneven" for the sixteenth trial. For if there is a system it seems to prefer uneven numbers, and if there is none, I can't make a mistake anyway (Fenk 1992). But how if the "series" that had produced uneven has the minimal length of only one trial? I would again bet on "uneven". And if I knew that on a certain day in a certain place on the equator the highest temperature was 40° C, I would – if I had to guess in the absence of any additional knowledge – again guess a peak of 40°C for the day after or the day before. The only way I can see to justify such decisions is an application of Occam's razor, or a principle at least inspired by Occam's razor: Do without the assumption of a change as long as you can't make out any indication or reason for such an assumption!

Hardly anybody would talk about laws in the example with the fifteen dices, or in the case of a series of fifteen S1–S2 combinations in a conditioning experiment, and most of us wouldn't even talk about "relative frequency" in our one-trial "series" – despite an ideal "relative frequency" of 1 in the one-trial "series" and in the S1-S2 combinations in the conditioning experiment. But the examples reflect a principle as *simple* as *general*: Use the slightest indication and all your contextual knowledge to optimize your decision but bet on continuity as long as you see no reason to assume that a system might change its output-pattern; generalize the data available to unknown instances! "Laws", "probabilities", and "objective chances" are – beyond a purely mathematical world – nice names for such generalizations and projections, usually based on large numbers of observations. But there is no lower limit regarding the strength of a regularity or the number of data available that ceases the admissibility of this way of reasoning! I can't resist quoting Hempel (1968:117) when he admits that "no specific common lower bound" for the probability of an association between X and Y "can reasonably be imposed on all probabilistic explanation."

4. Evolutionary perspectives

In his commentary on Campbell (1987), Popper (1987) agrees with Campbell's view of the evolution of knowledge systems as a blind selective elimination process. I am not quite sure if this is fully compatible with his remark (p. 120) "that in some way or other all hypotheses *(H)* are psychologically prior to some observation *(O)*". And principles of theory assessment such as Occam's razor might guide a systematic and conscious selection of theories in ways being more efficient and faster than a blind evolutionary process. Any sort of anticipation and of explorative or "hypothesis-testing behavior" imputes regularities and patterns and is successfull only if its heuristics and strategies in turn follow such patterns. The selective pressure was, first of all, on the evolution of mechanisms and strategies for learning risks and chances. In our recent life anticipation plays double a role: still as the cognitive component of any practical decision, and in science as the hypothesis tested systematically in order to improve our knowledge.

Irrespective of whether or not the evolution of knowledge follows a blind selective process: Real progress in nomological science seems to come about relatively slowly (Laszlo 1998), most apparently if predictive success or prognostic performance is taken as the relevant criterion, and in part due to an again "relatively" slow improvement of the respective methods. "Relatively" slow as compared e.g. with "vague but perhaps persuasive forms of explanation in the social and behavioral sciences" and "metaphysical theories of human nature" (Laszlo 1972:389) that cannot claim predictive success. A nice parallel in the evolution of technical equipment: "Using functional and symbolic design features for Polynesian canoes", Rogers and Ehrlich (2008:1) could show "that natural selection apparently slows the evolution of functional structure, whereas symbolic designs differentiate more rapidly."

---

2 For cases of two conflicting assumptions both satisfying the above criteria, Goodman (1973:94) suggests deciding for the assumption with the "better entrenched" predicate, e.g. for "all emeralds are green" rather than "... are grue", where "grue" "applies to all things examined before t just in case they are green but to other things just in case they are blue". But this argument is at best relevant if we don't admit any contextual knowledge. Why should we, on the expense of the precision of our predictions, allow all the emeralds having a specified crystal lattice to be either green or blue or to change their "output", i.e. the spectrum of the light reflected?

Literature

Campbell, Donald T. 1987 "Evolutionary Epistemology", in: Gerard Radnitzky and W.W. Bartley (eds.), *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*, Chicago and La Salle: Open Court, 47 – 89.

Coombs, Clyde H. 1984 "Theory and Experiment in Psychology", in: Kurt Pawlik (ed.), *Fortschritte der Experimentalpsychologie*, Berlin – Heidelberg: Springer, 20 – 30.

Fenk, August 1997 "Representation and Iconicity", *Semiotica* 115, 3/4, 215 – 234.

Fenk, August 2000 "Dimensions of the evolution of knowledge systems", *Abstracts of the VIth Congress of the Austrian Philosophical Society*, June 1 – 4 in Linz.

Fenk, August 1992 "Ratiomorphe Entscheidungen in der Evolutionären Erkenntnistheorie", *Forum für Interdisziplinäre Forschung* 5(1), 33 – 40.

Fenk, August and Vanoucek, Josef 1992 "Zur Messung prognostischer Leistung", *Zeitschrift für experimentelle und angewandte Psychologie* 39(1), 18 – 55.

Feyerabend, Paul K. 1962 "Explanation, Reduction, and Empiricism", in: Herbert Feigl and Grover Maxwell (eds.), *Minnesota Studies in the Philosophy of Science* III, Minneapolis: University of Minnesota Press, 28 – 97.

Forster, Malcolm R. 2000 "Key Concepts in Model Selection: Performance and Generalizability", *Journal of Mathematical Psychology* 44(1), 205 – 231.

Goodman, Nelson [3]1973 *Fact, Fiction, and Forecast*, Indianapolis – New York: The Bobbs Merrill Company.

Grünwald, Peter 2000 "Model Selection Based on Minimum Description Length", *Journal of Mathematical Psychology* 44(1), 133 – 152.

Hempel, Carl G. 1962 "Deductive Nomological vs. Statistical Explanation", in: Herbert Feigl and Grover Maxwell (eds.), *Minnesota Studies in the Philosophy of Science* III. Minneapolis: University of Minnesota Press, 98 - 169.

Hempel, Carl G. 1968 "Maximal Specificity and Lawlikeness in Probabilistic Explanation", *Philosophy of Science* 35, 116 – 133.

Hempel, Carl G. and Oppenheim, P. 1948 "Studies in the Logic of Explanation", *Philosophy of Science* 15, 135 – 175.

Hoefer, Carl 2007 "The Third Way on Objective Probability: A Sceptic's Guide to Objective Chance", *Mind* 116 (463), 549 – 596.

Huber, Franz 2008 "Assessing Theories, Bayes Style", *Synthese* 161, 89 – 118.

Laszlo, Erwin 1972 "A General Systems Model of the Evolution of Science", *Scientia* 107, 379 – 395.

Laszlo, Erwin 1998 "Systems and societies: The logic of sociocultural evolution", in: Gabriel Altmann and Walter A. Koch (eds.), *Systems - New Paradigms for the Human Sciences,* Berlin – New York: Walter de Gruyter.

Lewis, David 1994 "Humean Supervenience Debugged", *Mind* 103(412), 473 - 490.

Mises, Richard von [4]1972 *Wahrscheinlichkeit, Statistik und Wahrheit*, Wien – New York: Springer.

Nagel, Ernest 1961 *The Structure of Science*, New York: Harcourt, Brace, and Company.

Popper, Karl R. [6]1976 *Logik der Forschung*, Tübingen: Mohr.

Popper, Karl R. 1987 "Campbell on the Evolutionary Theory of Knowledge", in: Gerard Radnitzky and W.W. Bartley (eds.) *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*. Chicago and La Salle: Open Court, 115 – 120.

Rogers, Deborah S. and Ehrlich, Paul R. 2008 "Natural selection and cultural rates of change", PNAS Early Edition, 1 – 5.