AGORA —Papeles de Filosofía— (1998), 17/2: 25-38 ISSN 0211-6642

SCIENTIFIC PROBLEMS AND THEIR ROLE IN THE EVALUATION OF SCIENCE

Wolfgang Balzer
University of Munich

Resumen

La evaluación de los logros y problemas científicos se ha convertido en una tarea cada vez más importante. Generalmente, se afirma que esta tarea es política y que es casi imposible encontrar criterios puramente científicos para la evaluación de la ciencia. Yo sostengo que este punto de vista no es correcto. El argumento consiste realmente en formular un primer grupo de criterios para la clasificación de los problemas científicos (dentro de una disciplina y a través de diferentes disciplinas). Esto se lleva a cabo utilizando ciertos rasgos distintivos centrales que se encuentran en el modelo estructuralista de la ciencia. La mayor parte del artículo está dedicado al desarrollo y explicación de estos criterios y de la noción de los problemas científicos. En primer lugar, saco la consecuencia de que la filosofía de la ciencia puede hacer contribuciones prácticas importantes, por ejemplo, para la política de la ciencia y, en segundo lugar, el punto de vista estructuralista de la ciencia se corrobora posteriormente por medio de esta lograda aplicación.

Palabras clave: Evaluación de la ciencia, valoración de la ciencia, estructuralismo, problemas científicos, redes de teoría.

Abstract

The evaluation of scientific problems and achievements has become an ever more important task. It is usually maintained that this task is a political one, and that is is almost impossible to find purely scientific criteria for the evaluation of sicience. I argue that this view is wrong. The argument consists in actually formulating a first set of criteria for the ranking of scientific problems (within one discipline and across different disciplines). This is achieved by using certain central features found in the structuralist model of science. The main part of the paper is devoted to the develpopment and explanation of these criteria and of the notion of scientific problems. I conclude that, first, the philosophy of science can make important *practical* contributions relevant for instance for science policy, and that, second, the structuralist view of sicience is further corroborated by this successful application.

Keywords: Science evaluation, science assessment, structuralism, scientific problems, theory-nets.

The evaluation of science in our days becomes increasingly important. Several European countries have initiated programs for the evaluation of their universities, the other members of the European Union are likely to follow. At the moment, these attempts are of mainly political nature, and eventhough the evaluations are mainly performed by scientists themselves, there is little scientific method in their way of procedure.

Aceptación: III/1999

Evaluation of scientific achievement is a highly complex matter, the actual scientific development is influenced by factors many of which are external to science. Social values play a role as well as political preferences and often innerscientific factors are of minor importance.

Besides the evaluation of universities and their performance, another established way of evaluating science is found in the various institutions for the funding of research projects. These institutions evaluate research proposals in order to rank them so that the restricted amount of money available will be given to those projects ranking highest. Both kinds of evaluation, that of the performance of existing scientific units and that of the quality of scientific reasearch proposals, look different at a first glance. On a closer look, however, it becomes clear that, as far as scientific achievement is concerned, both kinds of evaluations deal with the same kind of objects and therefore should proceed in a similar way. The only difference is that in evaluating existing units one looks at their merits and achievements in the past which at the time of the evaluation are already known, whereas in evaluating research proposals one looks to the future and does not yet know the results. It seems to me that the future directed evaluation of research proposals is more fundamental because if we have a method of doing such evaluation we can apply the same method to evaluate also past and present periods. We simply have to look at the proposals which were made at the beginning of a given period, and evaluate the period in the future directed way according to the quality of these proposals. If this kind of evaluation should yield results very different from what is obtained by other, post hoc, methods then this either indicates that the scientific development was strongly influenced by external influences or that the method of evaluating by proposals does not yield the right results. Assuming that evaluating proposals is the best that can be done on a purely scientific level, any diverging results obtained by other methods would only show that these other methods include non-scientific, in particular political, aspects. On the basis of this argument I will focus here on the future directed evaluation of research proposals.

As research proposals are of extreme diversity they are not themselves well suited as a unit of systematic study. A more natural unit for the systematic study and evaluation of research proposals is suggested by the well known work of Thomas Kuhn in whose analyses the notion of a *scientific problem* plays a major role. In fact, the notion of a problem is not only suited for inner-scientific considerations as shown by Kuhns work, it also is suited to discuss social and political aspects of evaluation.[1]) Society and politicians are primarily interested in practical applications. But the lack of, or absence of, some practical solution whose need is felt or whose possibility is seen, is usually called a problem.

Therefore it is natural to reduce the task of evaluating research proposals

[1] See W. Balzer, 1999: «Eine Rolle für Probleme in der Wissenschaftskinematik», *Proceedings of the 1997 GAP conference*, to appear.

to a ranking of the problems which such proposals are investigating and to whose solution they promise to contribute. The notion of a problem also has the advantage of allowing the incorporation of non-scientific criteria in the analysis without change of terminology.

The choice among comprehensive scientific programs raises fundamental questions up to the level of moral questions, as we see presently by the discussions about cloning or about changing human genetic material. So ultimately, a ranking of problems cannot be achieved by mere innerscientific criteria; social values, political preferences and political programs will play a role. The ranking of scientific problems therefore cannot be a purely scientific task. Ultimately, it is a political task. What can be done scientifically is to study the innerscientific criteria for ranking and to use them to obtain a ranking of scientific problems which is based solely on scientific reasons. Such a ranking may or may not coincide with a more mundane ranking performed at the political level. If it coincides, fine, if not, the scientific ranking may possibly be used to criticise the social and political preferences.

Here we have a major task of the philosophy or theory of science before us. The task is to find purely scientific criteria of ranking scientific problems so that at a given time these criteria can be used to propose a definite ranking of a given set of scientific problems and to give advice to the political discussion and decision. This is a demanding task, and we are still rather far from good solutions. However, the progressive development of the philosophy of science over the past 50 years and its increasing specialization, and —if the term is allowed— professionalization, suggest that the task is no longer futile.

My goal in this paper is to take some first steps towards establishing precise criteria for the ranking of scientific problems. As a part of this, I will offer three precise definitions of problem types.

In order to set the stage for a detailed analysis, let us look at the development of science. This is a historical and social process which —as any process— may be conceptualized as a sequence of states. At each historical period, the scientific subsystem of society is in a particular state. In the transition to the next period the state changes, and a new state is realized. Of course, as a special case the new state might be identical with the previous one.

Such a state of science is a very comprehensive entity, including scientists, their products and problems, but also all kinds of institutions in which scientific work is done as well as important connections to the other subsystems of society, in particular to politics. I will be interested here only in the scientists, their output, and in scientific problems.

In order to describe a state of science at a given time I will use structuralist ideas and notions which however will be extended and modified for the present purpose.[2]) According to the structuralist model, a state of science

[2] The standard exposition is W. Balzer, C. U. Moulines, J. D. Sneed, 1987: An Architectonic for Science, Dordrecht, but see also W. Balzer, B. Lauth, G. Zoubek, 1993: A Model for Science Kinematics, Studia Logica 52, 519-48, for further developments pertinent to the present paper.

during a given period is represented by a so-called *theory-holon*. A theory-holon consists of various theories which are connected in specific ways. For the present purpose, let me use a pragmatically enriched notion of a theory, so that a theory consists of a hypothesis, or better, a corresponding class of models M, a set I of intended systems for this hypothesis, a set D of data structures each of which belongs to a definite intended system, and a set G of scientists acknowledging and using these items:

$$T = (M, I, D, G, r)$$

The intended systems are real systems to which the scientists intend to apply the hypothesis. For each such real, intended system the theory contains one or more sets of data which are obtained from that system. An indexing relation r is needed in order to spell out from which intended system a given data set arises.

A theory-holon contains *many* such theories which are connected by links or by intertheoretic relations. While a link between two theories establishes a connection between relatively few concepts of the theory, an intertheoretic relation involves the full apparatus of both theories; it spells out how each of the components: models, intended systems and so an, of both theories, are related to each other. Among the links I admit so-called internal links which relate different models of the same theory; these correspond to what in the structuralist literature are called constraints. The models of a theory are set theoretic structures of a type specific for that theory. The substructures which are obtained by «cutting off» suitable parts of those structures are of particular importance. Admitting substructures with finite and even empty components allows to incorporate finite sets of data, and a distinction between theoretical and non-theoretical terms in the picture. In particular it is possible to represent finite sets of atomic sentences by means of one substructure.

The notion of a theory-holon does not contain that of problems. Scientific problems are related to theory-holons in the following way. A theory-holon may be used to describe what I will call a space of problems, namely all those problems which are possible in, or relative to, that holon. This approach draws on ideas from cognitive psychology, notably from Piaget. In cognitive psychology it is now common knowledge that each problem requires a certain background of intellectual achievements and intellectual structures. If a person does not have that background, she will not be able to «see» the problem, or to formulate it, not to speak of solving it. In science, this background for «seeing» and formulating scientific problems may be modelled by a theory-holon.

In scientific development there is a dialectical interplay between theory-holons and spaces of problems. At a given time, a theory-holon induces the set of all problems which are possible within that holon. From this set of all possible problems by interaction with other parts of society a small subset of «real» or «acknowledged» problems is chosen. Some of these get funded and solved and their solutions add new elements to the holon. So the theory-holon changes and a new holon including the new results emerges in the next period. Then the cycle starts again.

The space of problems for a theory-holon cannot be explicitly defined in terms of that holon, for it is chosen in interaction with the social and political environment. Moreover, as we shall see in a moment, some types of problems have a pragmatic residue which cannot be fully formalized. Nevertheless, a set of *all possible* problems[3]) within a given theory-holon can be specified if a further pragmatic notion is added to a theory-holon, namely the notion of known procedures. The addition of the notion of possible problems therefore yields an important extension of the notion of a theory-holon. In the following I will assume that a state of science consists of a theory-holon and a space of problems for that holon which is chosen according to social or other criteria. The scientific development then is modelled by a sequence of changing states. The dynamics of such a development is driven by the interplay of formulating problems on the basis of acknowledged theories, and of solving them and adding the results to the stock of knowledge.

In this frame which I could only sketch here, we may try to construct a typology of scientific problems. At the present, first stage of research, I cannot claim to have a complete typology. But I can describe in detail three important types of scientific problems which seem to cover quite a substantial part of the array of problems which we find in the history of science.

The first type of problems I call *problems of data-fit*. Such a problem consists in finding out whether a given set of data from an intended system fits with the theory's hypothesis. When the hypothesis is represented by a class of models, the problem in its most explicit form is to find out whether for a given model and a given set of data, these two things fit together according to some standard of approximation which usually also is set by the theory but which I suppressed in the description of theory-holons. On a very coarse level we may introduce three values for fit: positive +, negative -, and unknown o. A problem p of data-fit in a theory T then simply is given by the three items: a model m, a set of data z, and the value of fit x, where x is one of +, -, o.

$$p = (m, z, x),\ m \in M,\ z \in D,\ x \in \{+, -, o\}.$$

Usually the positive case, when fit has been determined and turned out to be positive, is not called a problem; at best it is called a solved problem. but this is a matter of convention and convenience. The other two cases, when fit is negative and, most importantly, when it is unknown, certainly deserve the label.

[3] Note that this does not capture problems which arise from new hypotheses or data which are not yet contained in the holon. Such problems always are «possible», but are not included in the notion discussed here.

A second type of problems I call problems of *conceptual compatibility*. Such a problem obtains when two links and/or intertheoretic relations do not form a closed diagram. By a closed diagram I mean a triangle or a rectangle the edges of which represent models of different theories and the sides of which represent links or intertheoretic relations. The two most important cases are the following. First, consider a triangle with three models x, y, z of three different theories T_1, T_2, T_3 of the form

$$\begin{tikzcd} T_1\ x \arrow[r, dash, "\text{link}"] \arrow[d, dash, "\text{ir}"] & z\ T_2 \\ T_3\ y & \end{tikzcd}$$

where the vertical line represents an intertheoretic relation ir and the horizontal line a link. This means that there is a link between the theories T_1 of x and T_2 of z, and actually some concepts in the models x and z are connected by that link. Moreover, there is an intertheoretic relation between the theories T_1 of x and T_3 of y, and x and y are related accordingly. One would expect then that the intertheoretic relation is «tight» enough so that it allows a kind of «translation» of the link between x and z into a link between y and z. The absence of that link, i. e. the missing diagonal line from y to z indicates that there is a conceptual problem. The diagram is not closed.

The second case is a rectangle formed by four models of four different theories.

$$\begin{tikzcd} T_1\ x \arrow[r, dash, "\text{link}"] \arrow[d, dash, "\text{ir}"] & x'\ T_2 \arrow[d, dash, "\text{ir'}"] \\ T_3\ y & y'\ T_4 \end{tikzcd}$$

The two upper models are linked with each other, and each of them is vertically connected to a model from another theory by means of an intertheoretic relation. One would then expect that the two lower models also can be linked in some way. The absence of such a link, or the non-closedness of the diagram indicates a problem of conceptual compatibility.

In both cases one expects some kind of compatibility which would be expressed by the missing lines. From the second case we may easily derive a third case in which there are two links and one intertheoretic relation, while the second intertheoretic relation is missing.

The third kinu of scientific problems I want to consider are *practical problems* which however are of scientific nature. These problems consist in finding some *procedure* which produces an intended effect. Typical examples of procedures are experimental or measuring devices, like procedures to set radioactive markers on chemical substances, to chemically change a strand of DNS, to land a spaceship on the moon, or to move a body without using external force. A procedure in general may be described as a rule to produce certain *initial conditions* from which, by means of known theories, an intended effect follows by the laws of these theories. The notion of a rule has its known difficulties and it seems hopeless to achieve a set theoretic characterization, in particular in the present context where the rules under consideration are rules for the construction of artefacts. However, in order to get at least some closer connection with structuralist terminology we may use the following representation. We may identify a rule with all the results of its successful application.[4]) Thus for instance a rule for building a certain machine is identified with the set of all machines that *can be* built by applying that rule. In the present context of practical problems a rule is applied in order to produce certain initial conditions. Identifying such a rule with is possible results means that the rule is represented by the set of all initial conditions of a certain kind that can be realized with its help. In a certain sense the rule is replaced by its outcome. This certainly yields less than a full representation of a rule, but much more than a purely informal account.

Applying these ideas, I propose the following characterization of a practical (scientific) problem. A *practical problem* is given by a pair (INI, EFF) consiting some *initial conditions* INI and an *intended effect* EFF such that there are theories $T_1, \dots, T_n$ and models $m_1, \dots, m_n$ of these theories such that the following holds. Whenever the initial conditions are *distributedly satisfied* in the models then the effect will be *conditionally satisfied* in one of them. By *distributed* satisfaction I mean that each initial condition is satisfied in at least one model and that no model is redundant in this respect, and by *conditional* satisfaction I mean that the effect is satisfied in one of the models only if each of the other models satisfies «its» part of the initial conditions. We may imagine a real procedure as a real and complex process whose different aspects are simultanously or jointly modelled by different models of different theories, all of which are necessary in some sense in order to produce the intended effect.

It seems to me that with some good will also machines and static artefacts can be subsumed under this definition of a procedure. A procedure using a certain machine results in the production of something, of some intended effect, and the initial conditions are those that make the machine operating. A static artefact like a bridge can be seen as a limit case of a procedure, namely the procedure to transfer material bodies from one location to another one along a direct trajectory rather than making detours which otherwise would be necessitated by the environment.

[4] This strategy has proved useful in the context of measurement. See e. g. W. Balzer, 1985: *Theorie und Messung*, Springer.

In the following, I will assume that the problem space present in a state of science consists of problems of the three kinds just described.

We now may look at the evaluation of research proposals in a given historic period. The state of science in that period is described by a theory-holon which induces a corresponding space of problems. Each research proposal made in that period claims to contribute to the solution of one or several of the problems in the space of problems. The obvious, first move in evaluation now is to compare two proposals in terms of the *numbers* and the *importance* of the problems to whose solution they claim to contribute, as well as in terms of the *degree* in which they contribute to *complete* solutions. I use RP as a variable for the research proposals and p for problems. If d(RP, p) denotes the degree in which the research proposal RP contributes to the full solution of problem p and imp(p) the importance of problem p then the rank of a research proposal, rank(RP), may be defined by

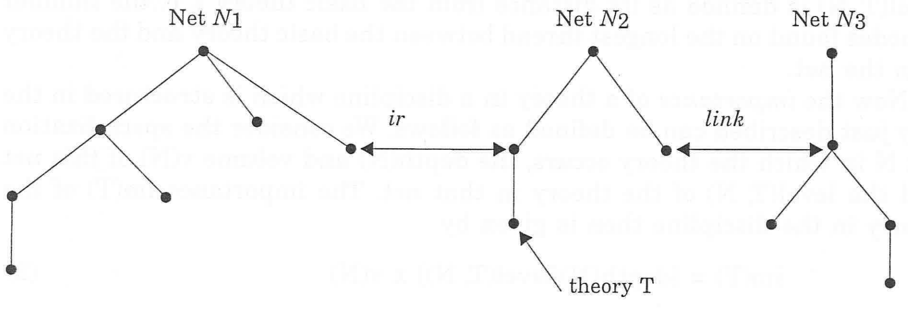
$$\mathrm{rank(RP)} = \Sigma p\ (\mathrm{d(RP,\ p)} \times \mathrm{imp(p)}) \tag{1}$$

where the sum is taken over all problems p to which the research proposal contributes.

The *degree* of contribution of a proposal to a problem cannot be further analyzed formally because it depends on the precision and the detail in which a proposal is formulated. In natural science and technology proposals often are very specific. They spell out exactly which problem is to be solved, and how. In such cases the degree of contribution of the proposal to the problem is simply equal to one. On the other extreme, proposals in the humanities are often not centered on a particular problem, and so it is difficult to estimate their degrees of contribution. However, the kind of ranking we are considering is a matter of approximation —as any other applied scientific analysis. With some idealization we may assume that well informed scientists even in the softer disciplines will be able to allocate problems and degrees of contribution to the proposals made.

More can be said about the other ingredient of the above formula (1), the *importance* of a problem. Let us turn first to the situation *within* one discipline. To this end I will assume that the theories in a discipline form a definite, comprehensive structure which is stronger than that of a theory-holon. This assumption at the moment is conjectural and there is no space to argue for it. This structure is best explained by means of the following picture.

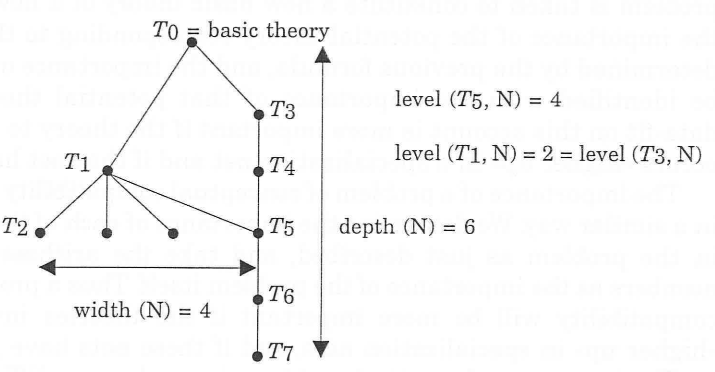
Figure 1



The theories in a discipline form various specialization nets. In each specialization net the theories (the nodes in the picture) are arranged in a tree-like way with exactly one so-called basic theory on the top of the tree. Each such tree is called a *specialization net*. Theories from different specialization nets may be connected by links and intertheoretic relations. Among the intertheoretic relations those of reduction and approximative reduction are salient. As has been shown elsewhere[5]) a reduction relation among the basic theories of two specialization nets automatically extends to the full nets, and a similar situation is expected for approximative reduction —though this has not been proved.

For each specialization net N we can define its *depth*, depth(N), and its *width*, width(N). The depth is the number of nodes occurring in its longest thread, and the width is the maximal number of nodes which occur as direct specializations of one common node somewhere in the net, see Figure 2.

Figure 2



[5] W. Balzer, J. D. Sneed, 1977/78: «Generalized Net Structures in Empirical Theories», *Studia Logica 36*, 195-210 and 37, 167-94.

The product of depth and width of a net I call its volume, v(N), v(N) = depth(N) x width(N). For each theory T in a specialization net N, its *level*, level(T, N) is defined as its distance from the basic theory, i. e. the number of nodes found on the longest thread between the basic theory and the theory T in the net.

Now the *importance* of a theory in a discipline which is structured in the way just described can be defined as follows. We consider the specialization net N in which the theory occurs, the depth(N) and volume v(N) of this net and the level(T, N) of the theory in that net. The importance im(T) of the theory in the discipline then is given by

$$\mathrm{im}(T) = [\mathrm{depth}(N)/\mathrm{level}(T, N)] \times v(N) \tag{2}$$

Thus a theory is more important if it occurs «higher up» in a specialization net, and if this net has greater volume. This is a first, admittedly rough proposal which perhaps needs numerical fine tuning. It is not difficult to normalize these numbers but no use will be made here of normalized versions.

The importance of a theory immediately yields the importance of a problem of data-fit. For such a problem by definition consists of a hypothesis and a set of data and thus corresponds to the two main parts of a theory. We therefore can regard a problem of data-fit as a kind of truncated, potential theory which can be inserted as a node in the theory-holon which describes the state of a discipline. In order to determine the importance of a problem of data-fit, we then have to determine the importance of this corresponding potential theory in the way described earlier. Eventhough this theory or node is only potential and initially need not occur in any of the nets, it is formally possible to determine its location, i. e. the right place where it should be inserted in the picture. To do this, one has to find out whether the hypothesis of the problem specializes one of the basic theories and to determine its level in the net of that basic theory. If no such specialization net is found, the problem is taken to constitute a new basic theory of a new net. In any case, the importance of the potential theory corresponding to the problem can be determined by the previous formula, and the importance of the problem may be identified with the importance of that potential theory. A problem of data-fit on this account is more important if the theory to which it is related occurs «higher up» in a specialization net and if that net has greater volume.

The importance of a problem of conceptual compatibility can be determined in a similar way. We determine the importance of each of the theories involved in the problem as just described, and take the arithmetic mean of these numbers as the importance of the problem itself. Thus a problem of conceptual compatibility will be more important if the theories involved in it occur «higher up» in specialization nets, and if these nets have greater volumes.

The treatment of practical problems is much more difficult. In reality, the weights attached to different practical problems vary enormously due to different political attitudes. These weights are determined mainly on non-scientific grounds. As I am dealing here only with scientifically based evaluation, these social and political weights must be ignored. It has to be emphasized that this may lead to rankings very different from those obtained on a social or political level.

Besides these social and political aspects there are two features of the importance of practical problems which can be evaluated on purely scientific grounds. First, the solution of a practical problem, that is, the invention of a new procedure, yields new successful applications for the theories used in the procedure. This follows from the way I defined the notion of a practical problem. What I called the intended effect of a procedure has to follow from the fact that one or several theories successfully apply to the situation. We can count, or rather estimate, the number n of successful applications of a theory used in the solution of a practical problem, and sum up the fractions 1/n for all theories involved in the problem. The idea is that if a theory has already many successful applications then a new one obtained from the practical problem will be less important. This yields a first part of the importance of the problem which may be called its *confirmatory importance* because successful applications add to the confirmation of their theories.

The second part of a practical problem's importance consists, broadly speaking, in the utility of its solution for the solution of other scientific problems, and therefore may be called its scientifically practical part. The analysis now becomes increasingly complex because this kind of importance introduces some kind of recursion. In order to determine the scientifically practical importance of a practical problem we have to refer to other scientific problems which can be formulated during the period under consideration but which possibly are not yet introduced in the range of acknowledged problems of that period. I will not attempt to spell out a numerical formula for this kind of importance but restrict myself to an informal account. For the sake of easier reference let me pick out one of the other problems for which the practical problem under consideration may be useful, and call this other problem the new problem. The question is, in which way the practical problem at hand is useful for the solution of the new problem.

Three cases may be distinguished. First, the new problem may be a problem of data-fit or of conceptual compatibility. As both these types of problems are treated and solved at the conceptual level it is difficult so see how a new procedure obtained by solving the practical problem could contribute to their solution. So we may dismiss these types of problems as irrelevant here.

In a second case, the new problem may be one which is not yet contained in the space of problems at the time considered. It could be that the application of the procedure provided by solving the practical problem yields new data, and that these new data generate a new problem of data-fit. As the production of new data is a contingent matter, further structural analysis does not seem appropriate. Rather, the probability of such events might be incorporated by means of a constant ß to be added to the practical problem's utility.

The third and most interesting case obtains when the new problem is itself a practical problem. In this case we have to estimate the utility of the practical problem at hand for the solution of another practical problem, or, in other words, the utility of the given procedure obtained from solving the practical problem under consideration for the invention of another, new procedure that solves another practical problem. Let us concentrate again on one of these other problems and call it the new problem. We have to check whether the given procedure contributes to the invention of a new procedure that solves the new practical problem. This can be done by looking at those procedures which are known in the period under consideration, and by checking whether some of these procedures can be combined with the given procedure to produce a solution of the new practical problem. The task is to find such procedures whose intended effects, when joined with the intended effect of the given procedure, yield a suitable set of initial conditions for the new problem. This involves reference to the theories which occur in the theory-holon but also to the stock of procedures known at the given time. Therefore, a formal analysis will be possible only if a further ingredient is added to the picture of a state of science, namely a set of known procedures. I leave this for future research.

The importance of the practical problem is thus recursively traced back to the importance of other practical problems, and the recursion stops when all its contributions to other practical problems have been checked. When a «terminal» practical problem is reached in this way, the second case previously considered applies, and the importance of the terminal problem is expressed in terms of the creation of new data.

The analysis up to now was based on the assumption that the theories in one discipline are connected so as to form a particular structure. If we look at the softer disciplines, especially at the humanities, it seems difficult to identify such a structure. Often, it seems that scientists in such disciplines do not aim at formulating precise hypotheses and test them against data. On the other hand, the structure of specialization nets, links and inter-theoretic relations in no way depends on formalization. On an informal level it is very general, and it only depends on the presence of singular and general statements. It seems to me that even in the softer disciplines such kind of structure can be found —though I cannot advance case studies in favour of this conjecture. Be that as it may, in the following I will assume that all disciplines have that structure.

On this assumption I can now turn to the final, most problematic, and most interesting part of evaluation, namely that of problems from different disciplines. This topic is usually treated with great discretion. The distribution of funds to the different disciplines in most cases is a matter of the historical growth of institutions, reaching back rather far into history. In Germany for instance, for reasons that do not matter here, physics has traditionally got the overwhelming part of the cake, in single universities as well as in national science foundations. The percentages in which the means are distributed usually are not a matter of bargaining and decision. Rather they are taken over from the past. Recently, formula based funding is spreading, but the main content of the formulas used just describes the distribution of funds as it was done previously. Changes in the distribution of funds typically occur in the way of diverting new, additional money to new, special projects which leads to an increase of the part of the discipline to which such projects belong to.

This system has worked quite well for several decades but with the increase of interdisciplinary research and interdisciplinary projects there is growing pressure to compare the scientific value of proposals and problems across different disciplines.

In the real world, the different sizes of the disciplines are mainly due to practical applications and related political considerations. At the social and political level it seems almost impossible to compare for instance disciplines in modern natural science with disciplines of more contemplative nature, like philosophy or theory of literature. However, we are not concerned here with social and political evaluation, and once we step down to the level of pure, scientific achievement, the picture begins to change.

In the spirit of the previous considerations, and assuming the discussed internal structure of the disciplines, a measure for comparing two disciplines with respect to scientific importance, merit, or achievement suggests itself. This measure consists of the sum of all volumes of the nets occurring in a discipline. For a discipline DIS with n specialization nets of volumes $v(N_1)$,..., $v(N_n)$ the measure m(DIS) of the discipline is given by

$$m(DIS) = \sum i = 1, \dots, n \; v(N_i)$$

In formula (2) above I introduced the importance of a theory relative to its discipline by [depth(N)/level(T, N)] x v(N), where N is the specialization net in which the theory occurs. If we replace the volume of the net by the measure m(DIS) of the whole discipline we obtain an expression for the importance of the theory as embedded in its discipline. Writing im(T, DIS) for the *importance* of theory T *in* the *discipline* DIS we thus get the formula

$$im(T, DIS) = [depth(N)/level(T, N)] \times m(DIS)$$

where N is the net in which T occurs. This formula may be used as described earlier to determine the importance of a given problem by finding the theories in terms of which the problem is stated, and by using their importance to calculate the importance of the problem.

In this way we obtain the importance of a problem in its discipline in terms of a number. Now two problems from two different disciplines may be ranked by comparing these numbers. On this account a problem is ranked higher if it occurs «higher up» in a specialization net, that is, if it is closer to a basic theory, and in this sense itself rather basic, and if its discipline has a greater measure. That is, the discipline comprises more specialization nets of greater volumes.

This is, I repeat, a first attempt which needs further study, in particular in terms of real life examples. Readers which are not familiar with the structuralist model of science perhaps may remain unconvinced, relying on their previous belief that evaluation is an exclusively practical, political matter. They are invited to take a look at structuralist case studies in order to see that the above considerations are based on experience with concrete theories and theory-nets.[6])

Still, one wonders how two problems from very different disciplines fare on that account. Well, let me very sketchy look at a case. Let us take the problem of what Aristotle means by «eudemonia» in his ethics and politics as a problem in philosophy, and the problem of cutting a string of DNS by chemical means at a certain distiguished position as a problem of genetics or biology. I would think that the philosophical problem gets quite some weight in its discipline. There is a vast literature on Aristotle, many authors have written on the notion of eudemonia. Moreover, the different treatments, or interpretations, of Aristotle do not form a homogenous mass. Some of them are more influential and are used and cited by others. This indicates something like the structure of a specialization net in the literature on Aristotle's ethics and politics. Also, there are discussions in which his metaphysics and logics are brought to bear on the problem, and there are comparisons to other accounts of ethics. These indicate something like a structure of links. Now counting depths and widths of the «nets» in the literature on Aristotle even in a very approximate way, it seems to me that the formulas I introduced will yield quite a substantial value for the given problem. The same holds for the genetics problem. There are some specialization nets in genetics[7]) and chemistry whose theories are used, and among these nets there are links and intertheoretic relations, so that the problem gets a respectable degree of importance.

The point here is not to obtain a definite estimate for the numbers expressing the importance of both problems in their disciplines. The point is that the philosophical problem does not seem to come out as hopelessly inferior. The reason for this is of course that external scientific criteria like the political and economic interests in genetic manipulations which make the genetic problem look so much more important at a first glance do not matter in the present, purely scientific evaluation. If scientific achievement is measured in terms of the internal structure and complexity of a discipline's output then the soft sciences have a chance.

---

[6] See *Architectonic*, note 2), and the bibliography of structuralism collected by Diederich, Ibarra and Mormann: W. Diederich, A. Ibarra, T. Mormann, «Bibliography of Structuralism 1971-1988», *Erkenntnis 30* (1989), pp. 387-407, and W. Diederich, A. Ibarra, T. Mormann: «Bibliography of Structuralism II, 1989-1994 and Additions», *Erkenntnis 41* (1994), pp. 403-418.

[7] See W. Balzer and C. M. Dawe, 1997: *Models for Genetics*, Frankfurt/Main, Peter Lang.