Explanation and Understanding through Scientific Models

Perspectives for a New Approach to Scientific Explanation

Richard David-Rus

Dissertation
an der Fakultät für Philosophie, Wissenschaftstheorie und Religionswissenschaft
der Ludwig-Maximilians-Universität
München

vorgelegt von
Richard David-Rus
aus Bukarest

München, 2010

Referent:	Prof. Dr. C. U. Moulines
Korreferent:	Prof. Dr. B. Lauth
Tag der mündlichen Prüfung:	10. Juli 2009

Dedicated to my parents and my wife

Acknowledgements

I would like to acknowledge my feelings of gratitude to several persons who have assisted me in different manners during my research. First of all my deep thanks to Prof. Dr. C.U. Moulines from the Seminar for Logic and Philosophy of Science at LMU for his constant support and steadfast advice he so kindly gave me during the elaboration of the thesis. I had the privilege to discuss with him some of the main ideas developed in my thesis during his doctoral colloquia. It is also due to his financial and technical support that this work was made possible.

My special thanks are also due to Prof. Dr. W. Vossenkuhl from the Department of Philosophy at LMU, for his encouragement and guidance especially during the first stages of my doctoral period and in the end for having accepted to be a member of the examination commission. I would also like to extend my feelings of gratitude to Prof. Dr B. Lauth for having accepted to be a member of the examination commission and for his kind and eloquent suggestions.

Furthermore my special thanks to Prof. Dr. Ilie Parvu from the Department of Philosophy of the University of Bucharest for his advice and guidance not only during the early stages of my doctoral studies but also at some key moments in the elaboration of this work. Some parts of this work were discussed and reconsidered following his suggestions.

I also want to express my sincere and special thanks to my parents and my wife for their spiritual and emotional support and encouragement all along the years of my doctoral work. It is to them that I dedicate this work.

Contents

Chapter one

Approaching the too many approaches

Anyone who intends to deal with the problem of explanation at the end of the sixth decade of the debate[1] may find himself in a rather uncomfortable position. By comparison to other topics he encounters a reticence of today's philosophers in addressing the topic in a direct way. This is due seemingly to a sort of fatigue or apparent exhaustion of the desire to deal with the subject.[2] On the other side, one can still find active working agendas concerning topics that integrate or make reference to certain aspects in the explanation topic. One such kind of agenda addresses topics related to certain philosophical aspects in particular areas of science and consequently touches on issues of explanation in that area. To mention only in passing such topics (which will be discussed in a greater detail later) one may recall: the causal mechanism type of explanation in biological sciences or the quest of explanatory virtues of simulations in different subfields of science.

Besides this, anyone who dares to approach the explanation topic in a straight manner nowadays is confronted with a spectrum of accounts comprising very different proposals that may lead to confusion or even inhibition, when getting through the topic. This chapter originates in an attempt to overcome this situation and to make the best out of it. In order to achieve this aim my strategy was to achieve a bird's-eye view of the multitude of proposals with the intention to extract some general characteristics, which may account at least partially for the divergence of approaches. At the same time I'll be looking for some particular ways to determine certain modalities of approaching the problem in a pertinent register. This procedure should allow me to evaluate further the strengths and weaknesses of the different sorts of approaches for a viable working agenda. As a consequence, I think, this will disclose the most plausible features that any future account would have. This endeavour could be also seen as an attempt to delineate broadly a kind of minimal program for an explanation account, given the actual context of the philosophy of science.

Nevertheless apart from the difficulties I mentioned in the beginning, there is optimism and hope that may be exercised as well. There are also certain advantages to work on the explanation problem in this period. In the meantime, the debate has cooled down considerably, and leading ideas were crystallized for certain positions together with different variations pertaining to them. By comparison to the previous decades, there is now a more generous offer of different perspectives to approach the topic, and the offer extends even more in regard to the means involved in working out the topic. There is a very rich environment for research in a good sense, despite possible feelings of "embarrassment" as Newton-Smith describes the today's attitude towards the subject.[3]

---

1 I'm following here the chronological setting from Salmon's referential text *Four Decades of Scientific Explanation*.

2 According to one of the philosophers (Berry Loewer in a personal communication), people got tired and bored by the subject so intensely discussed in the '70s and '80s.

3 In *A Companion to Philosophy of Science*.

Still why is there an embarrassment (or even a "scandal to philosophy of science" as the author calls it)? It is because, according to Newton-Smith, we still lack a unified account on explanation despite the existence of many pertinent in-depth studies. The need for such an account is pressing, given the widely accepted claim that the main task of science is to provide explanations. Besides this, claims in the realm of any other debates on different philosophical issues (as the one on scientific realism) need to be substantiated through reference to a concept of explanation.

From an historical perspective it is also a suitable time to draw some morals as to the fate of the debates in general, and how philosophical agendas evolve and influence other philosophical topics outside their main scope. In trying to react to Newton-Smith challenge, we'll acknowledge that promises were high for a general account on explanation. This also means that we have to keep a close eye on the past. From a more up-to-date perspective my conviction is that the explanation topic was not a dead-end; on the contrary, it is full of potentialities for advancing the philosophical insights into the nature of science and scientific activity. Moreover the aims of the philosophical approach to explanation have to be adjusted in the light of the results of the debate. In my thesis I will be arguing in a positive manner for such a type of adjustment. In the new framework of adjusted goals, inquiry into the explanation subject could bring a real contribution to the realization of a yet unfulfilled desire of the philosophers of science: to work intimately in conjunction with science.

In the following section I'll consider certain distinctions that will delineate the directions to advance a plausible approach. Distinctions have already been made in the literature from which I tried to cut across different accounts. These were often used to induce classifications over the accounts. A first way to induce such classifications as we find in different overview articles (e.g., those in encyclopedias) places the accounts according to a broader view that the authors have adopted as a working possibility. So we intend to regard the received view adherents or the ordinary-language analysts as approaching the explanation subject through their specific means. Another well-known classification used by Salmon identifies three basic conceptions: epistemic, ontic and modal, with the first type subdivided into inferential, information-theoretic and erotetic kinds[4].

In the flow of my argumentation I will make use of some of such previous classifications, which are widely accepted. But the distinctions I will be drawing do not aim so much to classify the existing accounts, but to suggest also possible directions that might help us to advance further solutions.

One can talk in terms of the intuitions behind the approaches – although the adequacy of an appeal to philosophical intuitions is nowadays heavily disputed. Viewing the task of the philosophical analysis in Carnap's terms as an explication of the explanation concept, these intuitions will be reflected in the choices made for the *explicatum* and especially in the modalities of its similarity to the *explicandum*. One could see a kind of core intuitions comprised in Hempel's model, which dominated the development of the subject during the first decades of the debate. Major subsequent accounts have developed on the split of the initial core: such as Friedman's or Kitcher's account exploiting the idea of explanation as a deductive argument or Salmon's focus on causality as *explicatum*. The fourth decade

[4] In *Four Decades of Scientific Explanation*, p.117 ff.

of the debate, in Salmon's own words one of the 'radical innovations', exhibits accounts totally and mutually exclusive in their assumed intuitions. Of course different classification as mentioned above, have to reflect the different intuitions.

Let us pursue with the distinctions:

Global versus local approaches

The first distinction I wish to point out is that between a local and a global kind of approach. This distinction is based on the way in which two different kinds of considerations (global and local) are to be seen as determinant for the scientific explanation and are therefore to be used in a conception of explanation. A kind of global-type of approach will consequently be one in which the global considerations are viewed as central. This does not mean that only explanations, which make direct appeal to the most general principles are proper explanations, but that the right criteria that determine an explanation are to be drawn properly from considerations at this level. Correspondingly, the same holds for a local view.

Let us see in greater detail how this distinction could be made more explicit and by what other means can we express it. We may look first at the uses encountered in scientific practice. Therefore we will usually say that the scientists solved a problem by global considerations if they make appeal to some general stipulations – e.g., such principles as those of conservation or invariance in physics – in contradistinction to the situation in which the contextual information extracted from the particular case at hand provides the solution. But it is not this sense that is intended by the philosophers who make explicit use of the terms. In the explanation debate, the terms acquired a particular meaning, while Friedman advocated a global approach to explanation. By drawing attention to the global aspects, he meant the aspects regarding the relation of the phenomenon to be explained with "the total set of accepted phenomena".

To shortly recall Friedman's account. He begins by stating that the usual *explanadum*-phenomena are expressed in the form of regularities constituting empirical generalizations or phenomenological or more special laws. They are explained through some other more abstract and general laws as is the case of Kepler's and Galilei's laws on motion that are explained by Newton's laws. The very essence of explanation in Friedman's conception is basically the reduction of a plurality of different law-like generalizations, which were previously accepted on independent grounds, to a more general law. The explanatory relation is ultimately expressed by Friedman only through set-theoretical relations; the two conditions providing the explanatory relation between two sentences being that the *explanadum* should belong to the *explanans'* consequences and the *explanans* should 'reduce' the *explanadum's* consequences. The notion of reduction from the last condition is expressed through an inequality between the cardinalities of the following sets: the K-partitions of the set reduced and its augmentation with the reducing sentence. This set-theoretical reconstruction of the explanatory relation can be viewed as making explicit the global aspect that determines the set of genuine explanations in science.

Kitcher's account provides us with another example of a *global* type of approach. As Friedman's one, it also explicates scientific explanation by connecting it to the systematization of knowledge. But in comparison to Friedman's account, unification is seen here to be realized through the repetitive application of a number of reasoning patterns to derive descriptions of different phenomena. The systematization of the corpus of knowledge that is generated by the set of argumentative patterns that best unifies the corpus will constitute the explanatory corpus, i.e., the set of explanatory arguments, which any valid explanation has to instantiate. A set of patterns will be qualified as most unifying by comparison to others, if it generates the greatest number of conclusions with a few stringent arguments. This way the global considerations related to sets of patterns determine primarily the explanatory virtue of the arguments.

In the view of the abovementioned authors, the *global* approach offers some major advantages for solving some difficult issues raised by the explanation problem. For Friedman the global approach gives an answer to the old argument against the existence of any genuine scientific explanation, argument that invokes the unexplanatoriness of the premises assumed. Besides, both authors see such an approach having the advantage to make clear the connection of explanation with understanding. A further benefit resides in the fact that it recuperates this way the objectivity of scientific understanding.

The more recent approaches, those developed by Bartelborth and Schurz (on which I'll touch in greater detail later) would make for good examples of global type of approaches. These authors pick on a sense directly from Kitcher's basic unification idea trying to articulate it in a more rigorous way. They do this either by adopting a well-articulated conception of scientific knowledge - the structuralist conception on science in Bartelborth's case - or by carefully formalizing the process of "assimilation" of the new phenomenon to be explained in the corpus of knowledge.

Examples of the local type of approach are those of Salmon and van Fraassen. In van Fraassen's case the emphasis on the contextuality of explanation reflects directly the local character of the approach. However this is not the way the abovementioned authors will use the term. For Kitcher and Friedman *local* will most probably characterize an approach that does not take into consideration the systematization aspects of scientific knowledge. As for Salmon, the attribute *local* is used for a type of explanation (not of an approach) that shows "how particular occurrences come about." He is viewing his causal/mechanistic explanation as a typical exemplification for this case. If an approach focusing on this type of explanation could be qualified as *local*, it is not clear from his discussion. Salmon is referring to types of explanations not types of approaches as Friedman. Of course for Friedman and Kitcher this type of explanation that focuses on singular occurrences was to be rendered only as a particular case, a limiting one in their approaches. But the issue has not been really addressed since their analysis builds on the assumption of *explanandum* being expressed through regularities.

In contradistinction, Salmon takes as paradigmatic for his analysis the explanation of singular events; in his view the identification of a cause implies taking into account the particular situation, in which the cause is acting. In his words: explanation is about opening the black boxes of nature and revealing the hidden inner workings. The very particular configuration of the situation to be explained is reflected in the network of particular causes at work that have to be captured by a valid explanation. A further specification that we should be aware of is that it is not that global considerations could not play a role in determining the explanatory relation. None of the two authors, Salmon and van Fraassen, will probably deny this fact. It is rather the fact that the global factors are not taken as the only considerations relevant in determining the explanatory quality; local factors are the ones that play the decisive role. For both authors, considerations related to unification or reduction of laws could guide the building of explanations; but if they are to be used and how they influence the building process, is ultimately determined by local factors.

To summarize: Salmon casts the difference global-local as characterizing two types of explanation (expl1 and expl2); this could be taken also to draw the difference between two types of approaches. But as we have seen, the meanings are not entirely identical when qualifying approaches or explanations. There is a difference in the way I'm using the terms local and global that takes Friedman's use as a starting point and becomes apparent through the fact that Salmon will not qualify van Fraassen's approach as dealing with local explanation. He regards it as rather neutral, since it could render either of the two types of explanations when additional constrains are attached.[5] By taking van Fraassen's approach as *local*, I plead implicitly for a broader meaning of the term *local*.

While keeping in mind these exemplifications I'll try to render more explicit the sense in which I read the distinction local-global by illuminating different facets through which it could be reflected.

A first plausible way that offers itself directly is to draw a separation in reference to the scientific entities invoked in the considerations that determine the explanatory relations. In the global type of approaches the authors are concerned with 'bigger' scientific structures. They concern primarily the aggregation of scientific entities, sets of statements or laws, or patterns, or models, making out the theories or corpuses of knowledge. Scientific knowledge is considered at the level of the overall structure and the way this organization is inducing explanatory relations. For Friedman it's about the relation of the *explanandum* with the entire set of the accepted phenomena; for Kitcher it is the membership of the pattern in the explanatory corpus E(K). Both accounts make reference to quantitative references invoking sets of statements or patterns and a number of axioms to which the statements are reduced or simply the number of patterns in a corpus. We can also claim that in global approaches there are usually proprieties emerging at the level of the bigger entities considered (special proprieties as unification or coherence) that bear directly on the explanatory power of these entities. I'll return to this point later in my thesis.

The local approaches do not imply such references; the structure of scientific knowledge is not particularly decisive in establishing the explanatory relation. For example, van Fraassen's semantic conception on theories is not brought to defend directly his account on explanation.

5 The point is made in the last part of the *Four Decades*, when Salmon seeks a compromise between the ontic (as his account) and the epistemic (as Kitcher's) approaches.

The classical Hempelian approach would have a somehow less clear positioning along this way of viewing the distinction or one that makes reference to laws as the main determinants of explanation. A natural move is to take theories as playing the central role in a global approach. As sets of statements or models systematized in a certain way, they can deliver the necessary determinant for the explanatory virtues. We could regard the universality of the laws as a sort of global feature in analogy with Friedman-Kitcher approach. And following the loose analogy we could search for a quantitative measure to render or to capture a feature, by pointing to the number of instantiations (or the degree of abstraction) of the law. Obviously we lack here the central idea – the systematization of the corpus of knowledge and the way it is achieved. The subsumption of facts under the laws or of more particular laws under more general ones through deductive arguments alone does not tell us much enough about the way we systematize knowledge.

There is a certain affinity between the DN-model and the global types of approaches. These authors are considered to develop their account by continuing the main lines of the Hempel's and the received view agenda. The idea of explanation as an argument is central to both Friedman and Kitcher; but the way they make use of it, by integrating it with the idea of knowledge systematization, underlines the radical difference from Hempel's model. There is also an affinity of the DN-model with Salmon's way of understanding explanation. Explanation1, the local type in Salmon's sense, directly captures one of Hempel's intended explanations: the causal type. Therefore Hempel's model can be viewed as containing *in nuce* the main ideas of both approaches without being reducible to either of them.

I mentioned briefly that van Fraassen's account can be regarded as a good example of a *local* type approach. The main reason for this would be the role pragmatics plays in determining the explanatoriness of a relation. The distinction pragmatics-non-pragmatics should, in my view, parallel in a good sense the distinction global-local. Pragmatics should play a central role in a local approach allowing in this way a better recuperation of the process of explanation as a part of the scientific process of inquiry. This way Salmon's meaning of the *local* is covered only partially. Though the causal-mechanism is not *per se* subject to pragmatic factors, picking the right causal network from the ideal explanatory text in Railton's view, which Salmon considers to complete the causalist view, is pragmatics dependent. So identifying the right causal mechanism and establishing in this way the adequacy of the explanation is influenced by the contextual factors of the investigation.

Another distinction that may be seen as recasting the wanted intuition behind the global-local distinction is the one used by Kitcher, i.e., between top-down and bottom up approaches. Salmon sees it as cutting along his explanation1 and explanation2. The bottom-up explanation appeals to the microstate of the facts targeted by the explanation, meanwhile the top-down appeals to more general principles and law-like generalizations. We encounter here a partial overlapping with the sense of global mentioned at the beginning of this section as being used by the scientists when drawing conclusions on global considerations. In an effort to synthesize and render the approaches complementary, Salmon sees both strategies: top-down and bottom–up, as being only different ways of reading the ideal explanatory text. Pragmatics determines in his view what way is to be picked in a particular situation. In as far as I am concerned, I cannot see everyone happy with this kind of compromise; there are passages of not unimportant details that are suppressed in order to carry out the overall match. One of such problem that becomes acute in this situation is the modality we should conceive how the ideal explanatory text is actually read in a global approach of explanation. This is in fact similar to the problem encountered by Kitcher when trying to account for causal explanation in his own conception.

In addition to the different ways of casting the meaning of the local-global distinction, I would like to make reference to another sense of the localism. The local character of the approach as specified above should also fall under this more general sense. The best statement of this sense can be found in Huggett's paper "*Local philosophies of science*". He calls *local* the approaches in philosophy of science if "philosophical problems are to be found and treated using the resources of more-or-less delineated scientific programs and not by trying to make science fit some prior vision". This sense should qualify the ones that I tried to pin down through the different distinctions. The modalities of casting the local character through the proposed distinctions are in a natural way qualified through Huggett's localism. This last claim is justified if we take into account the fact that the search for local considerations, which will determine the explanatory relations, is enhanced in the frame of a particular scientific program. On the other side global considerations tend usually to overrun the particular programs.

Static versus dynamic approaches

Another distinction that I would like to take into consideration is the one between static and a dynamic approach. It is not to be regarded as a distinction that can be drawn between different existing accounts on explanation. The existing accounts are more of the first sort. The distinction refers to a more general style of viewing the logico-philosophical reconstruction of scientific activity. The distinction is primarily between approaches that give a place, i.e., take more into consideration aspects of temporal evolution of the scientific structures *versus* the ones that ignore them. But how would temporal aspects be thought in the case of the explanation problem? Further on I will try to delineate the ways we can do this.

The static view is naturally seen as a legacy of the received view. It is not unusual to qualify as static not only the specific problem of explanation but the more general conception of science of the logical-empiricist view. As known, the historical orientation in the philosophy of science embodied the reaction against this aspect of the received view. Under the received view, the rejection of the pragmatics is the background that justifies the neglect of the temporal aspects. This should fall into the areas of psychological and sociological studies.

In the evolution of the problem of scientific explanation there are no obvious critical reactions to point to this aspect or rather to pick on it as the central point of their critique. We find scattered remarks and warnings only later in the debate (during the fifth decade) at such authors as Bunzl or

Sintonen. But someone could react by pointing to Kicher's account as one that addresses and integrates the temporal aspects of scientific knowledge into the explanation problem. This point seems right: Kitcher's conception can account for different corpuses of knowledge from different historical periods as determining the explanations accepted as valid in those periods. The account captures the dynamics at the macro level of scientific activity; a fact that could be seen as constrained directly by its global character. In my opinion, the sense of dynamics recuperated by Kitcher's account is only a specific one among many, and not the relevant one for an adequate solution. A proper dynamic account should unfold at a local level. To characterize it in a more general way, the main interest would be in describing how elements of scientific knowledge are modified or new elements are constructed in the process of providing an explanation for a phenomenon.

There are more ways available to unpack the above view. For example, Kitcher provides us criteria for comparing and selecting between corpuses of knowledge but he does not provide any clue as to how an explanation pattern is built. The solution in sight would be to provide rules as to how different elements of the pattern evolve, i.e., are chosen, modified or dropped in the course of searching for an explanation of a phenomenon. This will have of course to be integrated with the macro-constraints at the level of the corpuses. The task does not seem to be obvious if it is not even unrealizable. A relativization of the corpuses to more delimited fields of science will probably constitute a beneficial move.

In fact the dynamical aspects can be articulated in more diverse forms and Kitcher's one is only one among others. Variants that should be addressed are the ones that regard the way background knowledge is being modified when building an explanation. Bunzl, pointing to this issue, expresses it by requiring such accounts to be about "explanations as they arise in the context of building scientific theories"[6]. I'll tend to see this characterization extended to comprise also other types of scientific units, as for example, models. The dynamical aspects of explanation will become obvious if approached in the context of scientific modeling. Of course this is a particular angle of dealing with the problem, but a promising one as I'll argue further in my work.

Addressing the specific dynamics of the explanatory processes, one has to take into consideration the relations between different sorts of scientific representation. Aspects of the interplay between the representations should reflect the way scientists are looking for explanations, and the way explanations arise and are modified. Such types of accounts were in fact already developed by some authors as Hughes or Frigg and Hartmann, with reference to more specific registers of explanation i.e., explanation through models. I'll come to discuss these accounts and the virtues they exhibit in dealing with the explanation problem.

Making room for pragmatics should go hand in hand with the dynamic view advocated here. Contextual factors viewed at the local level shape in an essential way the process of scientific inquiry and the search for explanations. In a global approach pragmatics could be discarded to a certain degree. Kitcher's account is a pragmatics-free one but the modalities of comparing different corpuses of

6 In his book *The Context of Explanation*.

knowledge are not as much pragmatics-free as Kitcher wants to see them. Nevertheless it can be argued that the dynamics captured in Kitcher's account refers primarily to the units of scientific knowledge and not to explanations. Of course this claim has to be pursued in a more detailed way.

Explanation plays a central role in scientific inquiry and in particular in scientific discovery; it drives scientific inquiry and reflects in this way the process of knowledge expansion. This important aspect can be properly captured only in a dynamical account on scientific explanation. I'll touch also in a more detailed way on this aspect further in my work.

One may also rightly raise the question of the means that we could use to capture the dynamics. The dynamical aspects at the local level would be better exposed if we use specific frameworks. One such framework that proves to be suitable for this, as I'll argue in a later chapter, is the interrogative view on scientific inquiry. This will provide quite an efficient modality of modeling the dynamics of explanatory practices, as it will be shown.

Explanation as application *versus* explanation as construction

After a quick perusal of the previous discussion, I intend to advance a further distinction that might shed some light from another angle on the separation between the two types of views. The basic fact of one of the views can be put this way: it tends to conceive, to over-read explanation as application of a law or a theory (or other scientific entity taken as a central unit of scientific knowledge). I will call this sort of explanation, explanation as an application. The specificity of this approach is to be seen in the assumed (more or less) understanding of the application. First we should notice that there is no specialized, well-developed area of research in philosophy of science focused on the topic of application only. The term application is used usually in an informal way. We could say in general that an application of a law is an instantiation of it or that an application of a theory is a realization of it in some concrete situation – spelled out it would mean that some of the laws expressed in its sentences are instantiated in the situation under investigation. It could also appear that there are no further real issues related to this topic. Nevertheless many of the classical authors have in mind a sort of situation analogical to the logical-mathematical view. As in mathematical logic, an application will be thought along the lines of a sort of plugging values into a formula or formal schema. In the empirical sciences the situation is quite different from the mathematical one, and by using (even if only in a regulative way) the simplified version we distort the situation and situate us on an unfruitful track.

We could regard Hempel's and Kitcher's accounts as implementing directly this simplified version. In case of Kitcher's account, we can interpret his detailed description of the patterns (through filling instructions, classification rules etc.) as trying to deal with the limitation of the explanation as application view. I think we can find this attitude of approaching explanation as an application as a more constant and tacit assumption of the accounts in the classical positions of the debate. This last point is further illuminated if we look to the connection with another kind of distinction (to be discussed in more detail in the next section) – the one between theory-driven and non-theory driven accounts.

While we assume a broader view of application, we may regard explanatory processes as related to the notion of application in several ways. The important thing to stress and to take into consideration is the fact that applications in empirical sciences involve a series of scientific activities, which could be seen as parts of the application. Certainly it is more proper to see application as a construction process. There are several distinct activities involved in the application process such as decisions of what representation shall we use or what approximation and idealization should we build and adopt. From the perspective of the received view they will fall mostly into the realm of pragmatics and this could explain their neglect in the classical accounts of the debate.

How to characterize application is in my opinion not a one shot business. Making an attempt to delineate briefly some intended meaning I'll probably say that by using the notion of 'construction' I want to point to the different scientific activities involved in the process of explanation. Explanation as 'construction' takes seriously the idea that representation of phenomena must be constructed and it is through them that we may get an explanation. Such processes as idealization and approximation are part of this construction and of the explanatory practice.

The explanation as application bears also an analogy to what Cartwright in her critique of the received view on theories calls "the vending machine" view.[7] Theories are thought as one will "fed them an input in certain prescribed forms for the desired output" and after a while "it drops out the sought-for representation […] fully formed."

If there are chances to explicate the application concept in a more rigorous way, I think that the precise meaning should be picked under the local approaches on explanation as described in the previous paragraphs. But it is also highly plausible that we should regard it rather as an umbrella concept, which spans over different scientific activities, the explanatory activity being one among many. Referring to it here, my intention is to posit it as an opposite to the explanation as construction, in order to suggest the general direction of advancing the inquiry.

Explanation in a theory-driven perspective *versus* in a non-theory driven one

The 60s and 70s have witnessed strong debates concerning the nature of a scientific theory. This situation is to be seen also against the background of the reaction to the rather narrow logico-positivistic proposal regarding this topic. The rise of the semantic view offered a new alternative solution to the problem. But a more radical consequence of this reaction brought about distrust in the central role played by the notion of theory in the analysis of scientific enterprise. The concentration on getting in the first line a well articulated solution for the nature of a scientific theory, leaving the rest to follow from this or to be constructed around the notion of scientific theory, was gradually abandoned.

[7] In *Quantum Hamiltonians and the BCS model of superconductivity*.

Some aspects, which were thought as being secondary for the analysis of the scientific activity, gained much more importance; such as the experimental or the modeling activity and their products. Consequently, different topics, among them arguably explanation, gained (more or less) new valences in this new context.

The explanation topic rose at the status of a major subject in the philosophy of science in the heydays of the theory-oriented philosophy of science. Therefore it bears some of the legacy of that context. This could be seen also as one of the reasons why it is rejected in more recent philosophical agendas that assume a radical departure from any received view influence. Nevertheless we could read out influences of the theory-centric but also signs from the opposite attitude in today's approaches on explanation.

In the last and a half-decade one of the most active moves along this separation was undertaken by focusing on the role of models in scientific knowledge. Authors such as M. Morrison, Mary Morgan, Stephan Hartmann or Nancy Cartwright, are among the best known exponents of the modeling-oriented approach. In an effort to center the philosophical investigation on scientific models, they criticized the theory-driven view on models. What they mean by such a view on models is that models are seen as entities derived from theories and that their building, status and functioning depend in an essential way on theories. The authors engage in arguing for the autonomy of models, as for example in Morrison's way of making the point where models are autonomous agents in the production of scientific knowledge.

In an analogical way we can also argue for a more relaxed non-theory-centric kind of approach on the explanation topic. Even more, as I'll argue in the following chapters, we can take the position of considering models as a solution for the status of explanation-bearers, i.e., to consider the articulation of the accounts on explanation by making reference to models rather than theories or laws.

From an historical perspective one can distinguish between approaches on explanation that manifest tendencies towards a theory-driven view in a more or less obvious way. So, one will regard Hempel's account as theory-centric, since the concept of a theory plays a central role in its articulation. Kitcher's account integrates a theory-centric view too by advancing a proposal as to the nature of scientific theories as 'families of patterns'. In general it would appear as a normal tendency that global approaches should be more theory-centric biased. But different global accounts differ also in the role theories play in the articulation of the account on explanation. Some appeal in a straight way to the structure of theories as Kitcher or Bartelborth. But reference to more comprising units of knowledge such as corpuses of knowledge can bypass the appeal to theories, as in Schurz' case. Such a corpus of knowledge can be interpreted as a theory, but it is meant to accommodate various other forms of knowledge entities. In this sense, Schurz' account evades a strict theory-centric approach.

****

As I mentioned already the distinctions are not clear-cut. They are in a sense regulative, fixing some direction of plausible move. I'll try to circumvent some type of approach by reflecting it through different aspects and their opposites to specify better its intend.

As it has probably been noticed, there are affinities between the parts of different distinctions. I say this in the following sense: by adopting some direction to follow according to a distinction, certain options for the other distinctions will become more plausible to follow. As I mentioned, a global approach will tend to take the form of a theory-driven kind of approach. What I'll try to develop further in my work will be under the hope that a local, dynamic and non-theory-driven approach is the most plausible type to be pursued from the stance of the present landscape in the philosophy of science. (In fact it could be even claimed that the topics related to explanation problem in today's philosophy of science are already moving along this plausible lines sketched above).

My investigation in the next chapters will further take into more detailed consideration a particular way of articulating such an approach. I'll try to connect the explanation problem with a model-oriented view on scientific knowledge. I'll argue for the pertinence of following such an approach on the topic of scientific explanation.

Chapter 2

Explanation and models – bringing the subjects together

Salmon ends his overview essay written at the end of the 4^{th} decade with a section on the perspective to be encountered in the next decade – the fifth. Looking back we can engage in considerations on the fate of the debate in the last two decades following Salmon's essay. Such a quick overview will help us also to situate the main claims of my work.

As Salmon's career covered almost entirely the debate on explanation, and he engaged actively in it, his essay is a valuable document on the subject's evolution. Despite the fact that he acknowledges the particularity of his point of view reflected in his essay, his work offers an insightful account of the debate. Nevertheless the main attitude is marked by the author's position. It is straight to read his account as a liberation story, and the final conclusions expressed in his essay back up this interpretation. It is the liberation not only from the canon settled by logical-positivism, which raised the subject of scientific explanation at a major philosophical status; it is also the liberation from the philosophical misconception pre-dating the debate. This is the misconception that science cannot offer us explanations; this job was assigned to metaphysics according to the old conception. But of course the main liberation to read out from the final conclusions is related to the 'received view' type of approach. In this sense the received view and the kind of tools it deploys are rejected as viable ways to approach the explanation topic.

Though the story ends with the episode of the maturation of the new different positions emerged during the third and fourth decades, it still appears to be unsettled. In a good sense Salmon is under the nostalgia of the old and more unified approach. He looks therefore to the future state of the subject in the hope of an emerging consensus or reconciliation between the major positions.

However the consensus is not between all of the main types of approaches that he distinguishes: the modal, the epistemic and the ontic one; nor only between two of them. It is viewed to be only between some subspecies of the last two, and these are the unificationist and the causal/mechanical one. For the interrogativist approach – falling also under the epistemic type according to Salmon – he reserves a secondary role in the emerging consensus. This way Salmon proves to be quite particular in his hopes for the future agenda of the debate.

Taking his final conclusions as a departure point we can ask to what extent were his hopes realized during the following two decades of the debate. Furthermore the particularity of Salmon solution could be seen in the fact that at the moment of his writing, the hopes for a further advance were seen to lie in an approach that would recuperate also the pragmatic aspects of explanation. Sintonen's plea for a more involved pragmatics-oriented approach under an interrogative frame could be taken as a partisan position using the momentum at that time. But such a synthetic overview of the debate's landscape at the beginning of the 5th decade as Koertage[8] offers, gives us a less biased view that points also to the role of pragmatics.

In a recent paper[9] de Regt casts Salmon's complementarity thesis in terms of two kinds of complementary scientific understandings, which are seen as parts of a more complete understanding (called 'super-understanding'). Although from my point of view it is not clear how de Regt argues for the necessity of this particular explication, i.e., as being the only one that renders non-vacuous the complementarity thesis, his conclusions rightly restate my point that the "new consensus […] did not emerge."

We can also track the unfulfilment of Salmon's agenda for the 5th decade. No intensive activity towards this goal is seen, and no much consensus building takes the prime stage of the topic. Instead the majority of the worked out accounts pursued better articulations of the previous major ideas of the 'mature' positions. The relationship between the causal and the unificationist approach was still seen under the idea of reduction. Such well articulated idea as Bartelborth's and Schurz' are to be mentioned as constituting a continuation of the unificationist idea during the 5th decade; some authors, such as Sintonen or Kwajniewski, proposed further steps within the interrogative framework, meanwhile authors like Woodworth and Humphreys restated the causal approach in their accounts. There were of course hybrids and developments outside the strict ideas of the major positions. Hintikka and Halonen pursue their own agenda combining the interrogative means with the Hempelian idea of deductive inference. An exception that might fulfill Salmon's prediction could be seen in Strevens' recent kairetic account on explanation. He builds mainly within a causalist frame and ends up by appealing to the unificationist idea as an essential criterion for explanation. As I pointed out in my work, there is a specific way of articulating such a consensus, which differs from Salmon's idea as to what reconciliation should be. As it is elaborated in one of the chapters, Strevens' account is a sort of local one – his use of unification betrays the spirit of a pure unificationist approach.

To sum up briefly, the general tendencies, which one may notice to have occurred over the last two decades, are the following: an accentuated specialization placed on explanation in particular scientific areas, as for example, explanation in biological sciences or social sciences; a development of new accounts under the general core intuitions of old approaches, such as the accounts of Schurz, Bartelborth or Woodward, with a more clear articulation of the main concepts, attempting to overcome the problems raised by classical accounts – new *explicatum*-concepts were proposed such as a causal mechanism or the one that articulates the interventionist view of causality (Woodward); last but not least, there were attempts to propose synthetic accounts integrating elements from more than one of the major approaches such as the accounts of Weber or Strevens.

I am going to argue further for the pertinence of approaching the subject of explanation by considering it within the frame of the tendency in the actual philosophy of science that recuperates the

---

8 In her article *Explanation and its problems*.

9 In Wesley Salmon's *Complementarity Thesis: Causalism and Unificationism Reconciled?*, his interpretation is tailored after his agenda of promoting his conception of scientific understanding (which I'll discuss in detail in a later chapter).

models and modeling processes as central to scientific practice. I see such an approach as appealing selectively to the central ideas of the classical approaches without being subsumable totally under one such approach. These ideas will have to be articulated more explicitly in the modeling context, while a sort of synthesis will be achieved under precise conditions. This approach will ultimately induce the ways to select and clarify the concepts in particular areas of science, and a particular type of model that is taken under inquiry. In this sense the agenda is fulfilled in my view only in the frame of a particular scientific area and modality of scientific inquiry; no overall general valid account for the entire scientific practice could be sufficiently and clearly articulated. Nevertheless general ways of approaching explanation through models could be evaluated, and their pertinence, given the actual context, may be revealed. In the following sections I'll put forward some primary arguments in order to couple the approach on explanation with the one on modeling.

Explanation and models – stating the questions

Whenever someone points to this topic some direct questions arise. Some of the very first intriguing interrogations one may pose, they may look as follows:

1. Why should models be involved in the explanation problem?
2. What good news can the modeling view bring to the problem?
3. What were the major hindrances that blocked the "cross-hybridization" of these topics up to now? Why were the two topics held apart? What were the reasons to keep them separate?

Some of the first answers would probably look like the following:

Concerning the first question: the highly important role played by the models in scientific practice raises a kind of doubt: could the models be so insignificant for the explanation problem? Thus, this topic deserves a second thought at least in the actual context of renewed interest in the topic of modeling in science.

The focus on modeling is not quite a new fashion (even if a secondary one); in the sixties, there was a similar episode that emphasized the important role models can play in science. What is the difference from the actual revival of the modeling topic? The answer will surely point to the different context of basic assumptions and relevant problems for philosophy of science nowadays. In the 60s the agenda was still settled by the influent logical-positivistic orientation in philosophy of science and the reactions against it. At that time the main efforts were directed towards the search for a satisfactory conception of theories after (or in course of) the dismissal of the received view, and models were invoked and discussed having in mind a solution to this problem. The semantic view on theories provided such a solution taking into account the challenge to attribute models a more important place in the structure of science. On the other side there are critical voices claiming that this solution makes use of a too particular concept of model (the one used in mathematical logic) overlooking some of the very important features they exhibit.

As to the second question: the general benefit should be a more adequate view on scientific explanation. I think there are good arguments to conceive such a view as more practice oriented, i.e., taking into consideration the way scientific explanation is involved in scientific practice. Such a view, besides other items, requires an increased consideration for the modeling aspects of the scientific enterprise – as it will be argued further in the chapter. This attitude continues in a sense the reaction against the artificiality of the problems generated by the received view. This artificiality was generated mainly by the neglect of the richness of the practice of empirical sciences for the sake of a more direct application of formal means.

By pleading for a closer consideration of the scientific practices, I'm siding with a reaction that constitutes a much larger trend in the philosophy of science and advocates for this turn. This could be not much unheard news nowadays (as it was thirty years ago) but should be taken as a regulative impulse. A main line of separation for this could be drawn by taking into consideration, as the main guiding clues for a philosophical insight into the scientific enterprise not only the final products, the theories especially as exposed in the advanced textbooks, but also other items involved in the process of scientific inquiry. It points also to an effort to extend our analysis into the context of discovery as much as possible, bearing in mind that there cannot be a clear-cut distinction between the discovery and justification contexts.

An additional answer to the second question may point to the main incentive of pursuing such a project. This could be the disclosing of new resources from a too often ignored or misunderstood sub-field of philosophy of science as the one dealing with models. These resources could provide some new insights into 'old-fashioned' topics such as that of explanation.

In as far as the third question is concerned: the main reasons for the building of these hindrances are to be found in an attitude that acquired its best expression with the logical positivists. In their attempts to delineate as clearly as possible a meaningful kind of philosophical reflection from the meaningless way of philosophizing, they differentiate between two registers in which the scientific activities could be reflected upon: the context of discovery and the context of justification. The first could become the research object of psychology, the last of epistemology. The first one falls mainly in the realm of the pragmatics of science; the second could be reconstructed by syntactic-semantic means of investigation. The two subjects taken into consideration belong to the two different registers of investigation. The explanation problem falls traditionally in the justification context and therefore could be taken as an object of a logico-systematic kind of investigation. The modeling topic is, according to the traditional program, more adequately placed in the context of discovery, and therefore escapes such a logico-philosophical analysis. This difference has an effect on the orientation of the problems raised by the two topics, and the way the concepts are selected for each problem under discussion. (This attitude could provide also a reason why from a historical perspective the modeling topic played such a secondary role in the philosophy of science.)

For a more detailed articulation of this hindrance we should look into the assumptions and requirements set for the problems and solutions in contexts of the two topics, i.e.: a) in the context of the model topic, and b) in the context of the explanation topic.

Usually the subject of scientific models in the philosophy of science raised the problem of the roles models can play in science. The centrality of their roles and the relations between models and theories were the main subjects discussed under this topic. An influential conception with its main lines going back to Duhem's position on models in science, promoted a sort of depreciative view on models that denies any important contribution that models could offer for a philosophical insight into the scientific knowledge. Duhem's conception was spread mainly through the logical positivists' conception. On the other side, there is a view that emphasizes the centrality of the scientific enterprise in the activity of building and manipulating the models. In the next section of the chapter I'll review briefly the main ideas generated by the first view on models.

An historical detour: models and their relation to explanation

In this section I intend to trace the way the philosophers of science reflected upon models during the 20th century and the more or less explicit connection, expressed in some rare cases, with the topic of scientific explanation.

At the beginning of the century P. Duhem formulated, in his work on the structure and the aim of a theory, the main lines of a kind of "depreciative view" on models. Of course, it is necessary to contextualize his position so that we may be able to understand it in a proper way. Thus, in his book 'The Aim and Structure of Physical Theory' Duhem took into account the status and the development of the physical theories at the end of the nineteenth century, before the arrival of the two revolutionary theories of modern physics – relativity and quantum theory. But this would not be a major impediment for a pertinent reassertion of the main points of his position, even in the new historical context, only that it will give more weight to the adversative positions (as in fact happened).

Secondly, one of his main intentions was to distinguish and characterize some "national" styles of doing science; this intention is not "ideologically" uncommitted. He is engaged to plead for the French style of scientific thinking over against the English style. Some counter-arguments could make use of this commitment, but even allowing for their pertinence, there is enough room left to distillate a good position on models in science in its general lines. The later developments in the philosophy of science showed that the position could be held successfully by invoking further additional and different grounds.[10]

Duhem found the identity of a theory in the logical system it develops, the aim of which is to provide an economical classification of the phenomena, which it accounts for. The ways in which this is achieved is subject to different "cultural styles". Duhem distinguishes between two different kinds of

[10] I'm thinking here at the logical positivism view, which develops along the main Duhemian lines of thought, by making appeal to mathematical and logical means.

theoretical investigation: by modeling and by searching for highly abstract theories. These ways of doing science are generated by two kind of minds: an ample but weak mind (the English mind) and a narrow but profound (the French and the continental) mind. The first one relies on models as means of investigation in physical sciences. Models provide ways for drafting abstract principles, more understandable through materialization and concretization, by appealing more to the faculty of imagination than to that of reason. They are aids for the weak minds to help them work with the highly abstracted concepts of theory, making them more visualizable and palpable. In contradistinction, the second type of minds can deal better with the abstract constructs of the theory, which involve the reason's capacities directly.

In models, according to Duhem, the laws of the same group of phenomena are not "coordinated in a logical system" but are only represented. The logical requirements are more relaxed so they can freely proliferate and produce a disconcerting plurality of theoretical constructions. This aspect threatens the unity of the theoretical enterprise. What are then the positive aspects of using them? In addition to the already mentioned quality of representing in a more familiar way the abstract principles, which constitute the theory, they are very useful in applications, in engineering design, by facilitating the measuring and computational procedures. Here Duhem sees also one of the reasons for the spreading of this style of doing science on the continent. The only significant quality from the theoretical point of view, which could be invoked on behalf of the models, is the fruitful role they can play in the process of discovery. In this sense, the abandonment of the requirement of logical unity gives them more "suppleness and freedom" in the process of inquiry. But even so the models are not directly responsible, but rather the analogies. The fruitfulness in discovery is delivered by the use of analogies, which by their nature bring together two abstract systems – so that by using the one already known one can 'guess' the other unknown. Drawing on this idea, Duhem emphasizes therefore that even in the case of discoveries the role of abstract theories is central. Duhem concludes at the end of the chapter in which he discusses the status of models in science by claiming that the contribution of the models in science is rather meager in comparison with "the opulent conquest of abstract theories."[11]

As to the explanatory power of models, it doesn't seem proper to claim it for at least two main reasons. First, such a claim would not be possible as a whole in a Duhemian conception of science, since Duhem denies the possibility of science to provide explanations. Explanation understood as the unveiling of hidden ultimate realities beyond the appearances (the phenomena) is, according to Duhem, a metaphysical affair. Secondly, Duhem is suggesting[12] that such a kind of explanation (admitting that it can take place) would appear strange and inadequate given the fact that the models will unveil the reality of a realm in terms of another – so this would appear as an inadequate kind of explanation.

As for the status of models, Duhem's work lays down the main points of a kind of depreciative view on models. Models have only a secondary status, as they are provisional means deployed in the process of inquiry. The main functions they can fulfill are to provide a more understandable way of

---

[11] See *The Aim and Structure of Physical Theory*, p. 99.

[12] *Idem*, p. 75.

presenting a theory, to enhance this way the application of a theory, but also to act as heuristic means in the process of scientific discovery. There are also important dangers involved, of which the scientist and the philosopher should be aware. One such main danger is that this sort of thinking may threaten the logical rigor of the scientific enterprise. So, Duhem urges us never to take them too seriously in our scientific theorizations and cease our search for highly abstract theories in science.

The boldest early advocate of the central importance of models in modern science was Campbell in his book *Physics. The Elements*. His position is also to be viewed in the context of the advancements made by physical sciences at the beginning of the 20th century. He argues against Duhem's conception that reduces the status of the models to simple aids. In Campbell's view, the conception of models as aids stems from a bad reductive conception on theories, which identifies them only with logical systems, as in Duhem's conception. Campbell conceives theories as being constituted of two distinct sets of propositions, which he calls hypothesis and dictionary. The first set contains statements about ideas that are specific to the theory; the second set contains statements about ideas that link the theory's ideas with other ideas outside the theory. Analogies are essential components of theories providing the difference between an artificial, gratuitous calculus (an uninterpreted calculus, we would say using the logic terminology) and a scientific theory, which should exhibit an analogy with the real phenomena. The propositions of the theory's hypothesis must be analogous to certain laws (understood by Campbell as phenomenological laws). The analogy provides the value of the theory. The hypothesis gives real meaning to the theory by involving analogy. The dictionary is suggested by the analogy and makes use of it.

Analogies are fundamental for the development of theories according to Campbell, in opposition to Duhem's view. It is to be noticed that a conception that pleads for the importance of the models, could take up and promote the dynamical aspects of the scientific enterprise, central for the philosophical understanding of science, and as a consequence of the main role played by models in the dynamics of the theories. This remark shall become better articulated as we go further through our evaluation of the explanatory virtues of models.

The logical positivists were very "Duhemian" in spirit regarding the status of models in science. For Carnap they could have at best only a heuristic, didactic or aesthetic value. For Braithwaite, who pays greater attention to models, they are yet another interpretation of the calculus. The model differs from theory not in its logical order – both having the same formal structure caught in the calculus – but in its epistemological one. To be more precise, in Braithwaite's exposition, this means that in the model the logical prior terms determine the meaning and not the posterior consequences, i.e., the observational terms, as in the theory.[13] Models are therefore the most convenient way of thinking about the structure of the theory[14]. Nevertheless there are intrinsic dangers associated with the use of models, as that of identifying the theory with the model, or by thinking the theory in an 'as-if' way, or the danger of

---

13 See *Scientific Explanation*, p. 90.

14 See *Scientific Explanation*, p. 92.

transferring the logical necessity of some features of the model to the theory.[15] In another place,[16] Braithwaite provides us with an analysis of the ways models are used in theory extension and can provide novelty to the theory. He distinguishes four distinct possibilities according to the way new generalizations are introduced by means of new theoretical or observational terms. In two cases only, the assistance of models is not superfluous and it can "point" to extension "in the way which thinking of the calculus in isolation would not do." But even the most efficient way of 'pointing', by providing an argument by analogy, will render only a suggestion for any further extension of the theory.

Once the explanation-topic grew in importance, its relation with the modeling-topic appears to be taken into consideration by some authors. In Braithwaite's view we need models to make the concepts of the theory more understandable by connecting them to more familiar concepts. Braithwaite reflects on a possible model-centered account for scientific explanation as being a kind of "stage by stage" account: "Theory T_1 has to be understood by constructing a model for it, all of whose concepts are observable; theory T_2 can then be understood by constructing a model for it with the concepts of theory T_1 and so on". Therefore starting with the familiarity of the observable terms as an ultimate level, the explanation can be subjected to subsequent hierarchization. This may suggest a kind of conception of explanation as embedding[17] (as it was later developed explicitly and in a specific way in some semantic view accounts), specified in the case of the models by the reduction of the *explanandum* to the more familiar concepts of the *explanans*. In a more general statement and tributary to a syntactic view on theories, Braithwaite speaks of embedding laws in deductive systems.[18]

Nagel also dedicates some sections to the discussion of models in his well-known book *The Structure of Science*. He gives models a more important status in the economy of a theory, by making them the third component of a scientific theory besides the calculus and the correspondence rules. In a model the theory is fully interpreted. This does not mean that the theory is automatically linked to observation; it all depends on the character of the model. But, some models (the theoretical ones as Bohr's model of the atom) are important for logico-systematic reasons also, because the theory is necessarily expressed through them. In this sense, under the influence of the semantic notion of model, Nagel's view diverges from a strict Duhemian position in taking models only as some sort of dispensable scientific aids.

Similar to other authors, he stresses the familiarity of the concepts in models, and their use in providing a better understanding of a theory. But he also speaks of theories with unfamiliar models, which encounter resistance in their acceptance (as the models from quantum mechanics which can not be translated into more common, already known concepts). So, the familiarity of concepts is not a uniform characteristic of models.

---

15 *Idem*, p. 94.

16 See "*Models in the empirical sciences*", in *Readings in the philosophy of science*, ed. B.A. Broady.

17 Salmon comments also in the section of *Four Decades of Scientific Explanation* dedicated to the discussion of Braithwaite's book.

18 *Scientific Explanation*, p. 347.

Nagel also recognizes the important functions that models fulfill in scientific inquiry, especially the heuristic one. Models are used not only as a guide for setting up the fundamental assumptions of a theory but also to suggest extensions of the range of applications, once a theory is formulated (they are not dismissed as in the case of other authors) and further exploration of the main assumptions of the theory. They may even suggest new correspondence rules. These suggestions will point rather to the discovery context as the proper structure to place the issues regarding models.

But even if Nagel's position is more generous with the status of models in science, he does also mention the dangers posed by models for the theoretical aims of science (as did Duhem and Braithwaite, too.). Therefore, the main danger resides in the fact that features of the model could become misleading in considering the actual content of the theory. Of course this accusation is not so much different from the one raised by Braithwaite and the other authors. In as far as the relation between models and explanation is concerned Nagel does not seem to pay any attention to the subject.

After presenting several central ideas as provided by the accounts we reviewed above, it is now proper to draw some general conclusions. We have to remark from the beginning that we do not have to deal here with a unitary and well-articulated view on models,[19] even in the authors who belong to the same philosophical conception on science. Some authors discuss the notion of model under that of analogy; while for the late logico-positivists, one may notice a rather different position, due perhaps to the increased influence of the mathematical notion of model. The analogical models are for the latter only a special type of model, even though a most representative one.

A more unified perspective is to be found in the reflections regarding the roles models can play in the scientific enterprise. Most of these roles fall in the register of the context of discovery, but as already mentioned, some types of models can exhibit logico-systematic functions. This kind of function is properly investigated by rigorous means in the justification context. Nevertheless, the results of such an analysis are only secondary, without importance for the more general systematic structure of science. I think that here we could blame the depreciative view on models for this situation. The view induces strong inhibitions on any possible logico-systematic important consequence that models can reveal for the study of the structure of science. Moreover, under this view there are more characteristics assigned as main features of the models which bear negative consequences for the main goal of the scientific enterprise.

From the very beginning models are limited in their possibilities as scientific constructs. They are subject to different limitations affecting as such the ideal of the universality of scientific claims. They exhibit a temporal limitation, when viewed under their heuristic functionality. The scientists will dispense with them as soon as they get the new theory. Even when they remain functional under the new theory, they are limited in their range of application. Likewise the last point may be thus interpreted that limitations are to be found in the possibilities of the models to express a scientific result in its utmost generality. The limitation of the models in their logico-systematic capacity triggers in the abovementioned authors the warning against the danger of taking them more seriously. The main

[19] The observation was also made by Achinstein in *Concepts of Science*.

sources of these dangers are to be found in the lack of logical structure of the models. As Duhem has already emphasized, the laws appealed in the models are not coordinated in a logical system. This can have as a negative consequence, according to Duhem, a free proliferation of models, meaning a weakening of the systematic control of the theory.

Another type of accusation brought against the models was the one of (we could call it) ‘improperness’. A much too serious involvement of models could lead to improper results, which could pose significant dangers for the scientific enterprise. One may notice here that the dangers, as presented by different authors, could be put in two registers: one regarding the syntactical nature of the scientific constructs, and the other the semantic aspect. The ‘improperness’ at the syntactic level refers to the possibility of improper use of some terms generated by the appeal to the terminology of one theory for the investigation into the domain of another theory. The second refers to the improper involvement of the semantics of a theory to shed light on the semantic aspects of another theory (as Nagel and Duhem suggest). For Nagel this could lead to a wrong evaluation of the content of the new theory. The ‘improperness’ is also expressed through the confusion raised by the appeal to different ways of thinking, as Braithwaite emphasizes. In this way, the danger lies in confusing the rigorous way of thinking of a theory with the more relaxed as-if way of thinking of a model. It enhances further the danger of an unallowable transfer of necessity attributed to propriety in the model to the necessity of that propriety in the theory.

Now, in the context of the explanation problem, it is befitting to discuss Hempel’s reflection on models and explanation. In a positive sense and from today’s perspective the explanation problem is not reducible to Hempel’s appreciation of the problem but, interesting enough, even when the participants in the explanation-debate were not any more committed to the requirements of the Hempelian program, the models did not gain in importance in the search for solution. An answer to this situation can point to the latent legacy of the logical positivism; in this sense the fundamental standards of the received-view, which determine what kind of solution the explanation problem requires, are still valid in later approaches.

Hempel’s explicit reflection on explanation through models and analogies

Hempel is the author who presents us an explicit reflection on the explanatory powers of the models. In a section of his *Aspects of Scientific Explanation,* where he discusses the subject, he states to start with that in the empirical sciences, sometimes, explanatory accounts “are formulated in terms of a *model* of the phenomena to be explained or in terms of analogies between those phenomena and others that have been previous explored.”[20] He engages further in discussing the explanatory claims that can be invoked for analogies rather than for other sorts of models. Explanation by analogies could recall the

20 *Aspects of Scientific Explanation*, pp. 433-434.

particular position[21] that explanation in science involves a reduction to the familiar. Of course, this is an approach to explanation, which Hempel regards as fundamentally incorrect. In his further analysis of the analogical models, he points to the fact that an analogy implies a “nomic isomorphism” – a syntactic isomorphism between two corresponding set of laws (governing the behavior of things in the two domains). Moreover, to say that a system is an analogical model for another system is rather an elliptical claim according to Hempel, since only a subset of the set of laws, which govern the two domains, are isomorphic. The laws, which are isomorphic, are usually, as in the example of the mechanical models for atoms and molecules, few in number and limited in scope, so that several different models are used to represent different aspects of the modeled system. In the case of fruitful analogies, the isomorphic laws or theoretical principles are stated in mathematical terms, e.g., as the mathematical theory of wave motion, strong enough to allow deductions of further important laws for each domain.

So, the main benefit of using this isomorphism is that it can play an important role in the discovery of the laws of a new domain. Once we get the entire set of laws for a new domain, we can drop the analogy and use directly the new laws in our explanations. In Hempel’s own words: “for the systematical purposes of scientific explanation, reliance on analogies is thus inessential and can always be dispensed with.”[22]

The analogical models proved to be useful in other respect than for explanatory purposes. They provide the needed “intellectual economy” for scientific reasoning by involving old laws for new domains. So for example, such models proved to be useful in devising and expanding microstructure theories as the kinetic theory of heat. The numerous models from elementary particles in physics constitute a similar situation, but they also differ in some respects from pure analogical models; they constitute according to Hempel a sort of provisory theories, which are “put forward *until further notice*”.

In a similar way to the other authors mentioned earlier, Hempel also stresses the important quality of models in providing understanding. Models can “facilitate one’s grasp of a set of explanatory laws or theoretical principles for a new domain of inquiry”, contributing this way “to the pragmatic effectiveness of an explanation”.[23] Understanding being for Hempel a feature of explanation that belongs entirely to the realm of pragmatics, it follows that models contribute more in this register to the explanations. In this way, they evade a more systematic analysis of their explanatory pretensions. Following the same idea, Hempel points to the fact that models proved to be more useful in the context of discovery (than in the context of justification); while by itself an analogical model explains nothing, it can provide suggestions for further extensions of a theory, that means heuristic guidance.

At the end of the section, Hempel mentions also yet another type of model, the theoretical or mathematical model. He specifies as paradigmatic examples the mathematical models in social

[21] As expressed, for example, by Toulmin.
[22] *Aspects of Scientific Explanation*, p. 439.
[23] *Idem*, p. 441.

sciences; further on in the section he acknowledges that “such peculiarities” are to be found also in “the field of physical theorezing”. In comparison with the other types of models, it can be claimed that the last ones are a sort of theories. The difference from proper theories, especially the advanced theories in physics, lies in their limitation in the range of application and simplification of the assumptions they make use of. Otherwise, the theoretical models can fulfill the same functions in theoretical systematization, scientific explanation and prediction like the normal theories. The limitation in range is due to the employment of various idealizations and oversimplifications. This is also the main reason why Hempel sees them as having restricted explanatory powers: “a limited scope and only approximate validity within that scope may severely restrict the actual explanatory and predictive value of a theoretical model.”[24] It is worth to remark that Hempel does not deny their explanatory virtues, but sees them in the best case as a particular case of his deeper and more general account on explanation.

The bearers of explanation

The explanation-problem is of course not reducible to the Hempelian approach on the subject and the debate around it. It includes also the ulterior modifications and developments of the topic, which advance different sorts of requirements for the solution of the problem. But even so, if we take into account the most important positions in the debate, it may be claimed that, in the context of the explanation-problem, models were in general ignored by the majority of the participants as serious candidates for being the ‘bearers of explanations’. I’m using here the term ‘bearer of explanation’ to name the units of scientific knowledge, which are required by the explanation. They can be more or less explicitly articulated, ranging from laws and theories to more comprising forms as corpuses of knowledge (Kitcher) or fields of inquiry (Schafner).

A bold question arises in regard to the degree of involvement in the problem of explanation, i.e., of specification, and the role played by these units. I don’t think that this can be determined prior to the construction of an account on explanation, as for example through a discussion or a general argument that should claim that a satisfactory solution should be searched only under some well-determined kind of specification of the units of knowledge. I think that it is more probable the case that this involvement is only an implicit consequence of opting for and proposing a model of explanation under a view. I’ll try further to spell out this claim by exemplifying it under particular views.

A first such important position is the one for which a solution for the explanation-problem is to be built only by being supported by a good articulated theory about the structure of science. Consequently, such a strategy will single out some specific substructures or some relation among them,[25] as being directly involved or responsible for the explanatory activity. I think this view will promote more directly the attitude of explanation as an application. Typical examples of accounts

24 *Idem*, p. 447

25 As in the case of Friedman’s approach, where the relation of reduction between laws makes out the explanatory pattern.

falling under this position are Hempel's model or Bartelborth's proposal. Hempel's model is designed to fit the received view of the structure of science. Bartelborth's account presents itself as a direct application of the structuralist view on science for the problem of explanation.

An opposite view is the one that disregards any account of the structure of science as being important for a solution to the explanation-problem. Of course, there could be (or even it has to be) a kind of reference to some forms of knowledge on the base of which the explanation is given. But the analysis of explanation takes place by involving some background knowledge not necessary spelled out in well-determined forms. The forms in which the scientific knowledge is organized, are not considered to be essential or a precondition for the explication of explanation. Examples falling under this view are the accounts of van Fraassen, Kitcher or Salmon. The first two authors develop conceptions regarding the forms of organization of scientific knowledge. But in van Fraassen's case his semantic conception on scientific theories does not bear in any direct way on the articulation of his pragmatic model of explanation; while in Kitcher's case his view about the way scientific knowledge is organized seems to follow from the account on explanation rather than to determine this account.[26] Accordingly, we could distinguish some different strategies of building under this register.

For example, a strategy developed under this view is the one that takes the explanation mechanism as primitive for an account of scientific enterprise, and shows how it contributes to the generation of the units of knowledge. Kitcher's model, with its emphasis on the cognitive mechanism of explanation, constitutes the boldest example to be mentioned here. As I have already mentioned when presenting his account, the explanatory patterns build up families, which further form corpuses of knowledge**.** But the main problem with Kitcher's proposal is that there is a lack of clear identity criteria for such families, as one may accuse.

Another strategy under this view will be to renounce entirely any reference to units of knowledge. In Salmon's account no such reference is made when we point to the causal mechanism responsible for the *explanadum*. This is also the case with van Fraassen's pragmatic model. No attempt is made in the last conception to implement the interrogative mechanism into some forms of knowledge. Even so, in van Fraassen account the qualification of the explanation as scientific takes place by appealing to scientific information. But criteria for the scientificity of the information are not delivered.

The choice between the two views regarding the clarification of scientific knowledge could also be seen as rooted in an implicit option regarding the priority in the investigation of explanation in science *versus* the explanation in general and the relation between them. On the one side lies the project of reading the mechanism of explanation as it appears in science. The more ordinary form of explanation (the *folk explanation*) is that obtained by a relaxation of some of the conditions of the scientific explanation. For example, the universal laws from the DN-model could be relaxed to some folk-generalization to get the everyday explanation by subsumption of particular facts under some generalizations.

---

[26] Kitcher's position regarding the bearers will become clearer later in this section.

One may claim that there is a kind of neopositivistic perspective contained in this project. It builds on the assumption that science is the place where the explanatory mechanisms reach their full exemplar form and where we should focus our investigation for a solution to the explanation-problem.

The other implicit option assumes that the mechanism of explanation is not to be necessarily revealed in the context of scientific knowledge. The discovered mechanism is extended to the scientific context by providing several constraints in addition, typical for this kind of knowledge. In Kitcher's model, for example, the global constraints, as pointed earlier, are the ones responsible for the scientificity of the explanations.

Regarding the above discussed choice, my option is one which favors the importance of the scientific forms of knowledge for the explication of the concept of explanation. Without pleading in favor of a neopositivistism, I think that by investigating the ways science provides us explanations, we are in a better position to search for an answer to the question of what it means to have an explanation. The forms in which the scientific knowledge is organized, backed by the fact that we can have some well articulated accounts on them and the way they are used, constitute great advantages for the investigation of explanation. An inquiry into scientific explanation – the most elaborated form of explanatory activity – must also involve the ways the information is structured and used in scientific activities, as main clues for its investigation.

If we are now to ask what are the most typical invoked bearers of explanation, we will probably point to the laws and theories (perhaps also due to the dominance of the Hempelian model and the discussions around it). But the above claim is in need for more clarification. While laws, theoretical or phenomenological appear explicitly in the explanatory schema provided by some accounts as the one of Hempel or Friedman, there is no account to refer to theories as direct components of the mechanisms, even if they are claimed to be explanatory. Furthermore there might be, from other argumentative perspectives, good reasons to take them as units of knowledge on which the explanation is based. The further paragraphs will spell out these claims more clearly.

There is a first ambiguity regarding the distinction between laws and theories as bearers of explanation in Hempel's account. Hempel's definition of explanation, as one may gather from his 1948 paper, cites explicitly theories in the formal explication of the concept of explanation. Theories are defined there to differ from laws only by the presence of the existential quantifiers beside the universal one in their formal expression. The theory cited in an explanation is not expressed by a single sentence only, but by a conjunction of generalized (with both types of quantifiers) sentences. But Hempel's notion of a theory, as appealed in the explanatory schema, is not the one that it is usual articulated in the received view (as in Nagel's classical book, for example). It could be said that the theory (in Hempel's sense) cited in the explanation comprises rather fragments, i.e., some of the laws (relevant for the explanation) provided from the stock of laws of the received-view type of theory.

In the standard presentation of the received view – as expressed by such authors as Carnap, Nagel or Braithwaite – theories are in a position of being only indirect explanatory. Under the assumptions of the syntactic view that conceives theories as being axiomatizable sets of sentences, a theory would explain some facts if it contains between its theorems the laws cited in the explanation. So, the form of knowledge cited directly in the explanation is expressed through some of the laws of the theory, among them also a phenomenological one (in case of a factual *explanandum*), and not the whole theory. Most of the theoretical laws enter usually in the premises of the schema of a theoretical explanation, i.e., in explaining further more specific laws. The claim that the theory explains the facts is therefore a sort of an elliptical claim in the sense that the spelling out of an explanation makes recourse only to some parts of the theory, articulated in adequate ways.

But even if we resume ourselves to Hempel's explanatory schema only, the case for the theories as the real bearers of explanation could still be built up. The explanatory schema cites more information than the one contained in the laws, and that is the information contained in the initial conditions. A question arises here, where should one find the kind of knowledge appealed to in order to get these conditions? If we are to take into account the difference between the theoretical part of a domain and the experimental one, the balance seems to incline more on the side of the latter. This does not mean that the experimenter's knowledge cannot be made explicit or be reflected in an account on scientific units of knowledge. In fact, the explanatory schema does not make reference to such distinctions in its articulation, and neither do the accounts on the structures of science. This kind of knowledge is to be reflected at least partially in some elements of the structures of science.

In case of the standard "received view" conception of theories, some of the elements, which play a central role in this process, are to be found under the correspondence rules of the theory. These rules tell us how we should interpret our abstract symbols from the theory's calculus, providing this way empirical content for the formulas. To get the initial conditions, it would mean to provide further the necessary description of the factual conditions, responsible for the instantiation of the law-schemata. It means implicitly the capacity to identify the referents for some of the concepts of the theory, expressed by symbols in the theoretical calculus. In a more mundane way of putting things, the process of providing the initial conditions is caught into the knowledge of several experimental techniques and protocols for the identification and measurement of the quantities of the theory. So taking into account the necessary reference made to this kind of knowledge, the bearers of explanation in the DN-model are the theories or at least will require a more consistent portion of them comprising besides laws, the guiding knowledge for determining the initial conditions. I draw on the received-view conception as a possible register to represent the structure of a theory, but the above conclusion holds also for the more particular Hempelian position[27] regarding the components of a scientific theory – as being constituted of internal and bridge principles. The difference from the received view is not relevant for the matter under discussion. Instead of the correspondence rules, the components of the theory required in Hempel's conception for capturing the empirical content are the bridge principles.

Now, the DN-type of explanatory schema may require an appeal to more than one theory for an explanation – a situation that is not unusual in explanatory practice. The cited laws in the first premise

[27] As it is presented in *On the Structure of Scientific Theories* or in *On the "Standard Conception" of Scientific Theories.*

of the explanatory argument could belong to more than a single theory, and the same can be said about the knowledge necessary for the initial conditions. So, certain bigger units, as theory-nets or fields of scientific investigation, as Schaffner uses in his account,[28] could be claimed as bearers. One may claim that this kind of explanation is more adequate for accounting for explanations in interdisciplinary fields of scientific investigation.

The above way of making recourse to the received view or Hempel's conception is a very reductive way of accounting for all the richness of the experimental knowledge by way of pushing it indistinctly under a single concept – that of a rule of correspondence or bridge-principle. This kind of knowledge has to be reflected in some way in the explanatory schema. The request for a more articulate specification of knowledge in this case, and its involvement in explanations is therefore entirely justified. Kitcher is the author who provides us, more than any other participant in the debate, with such a good articulated specification of that knowledge. The fine cognitive structure of the explanatory patterns offers a sufficient good answer to the question of how the information appealed to in an explanation is organized. But, the case for some bigger units than the patterns as bearers of explanation could be rendered even stronger in Kitcher's case. In a sense, Kitcher makes explicit exactly what the above paragraphs on Hempel's model pointed to: the need to get a sufficiently complete unit of knowledge responsible for the explanatory virtues manifested in the scientific enterprise. This is addressed in his account in a way in which the explanations are determined by a more comprising unit of knowledge. As he puts it: "successful explanations earn that title because they belong to a set of explanations, the explanatory store." So, the systematized corpuses of knowledge are the candidates for the 'ultimate' bearers rather than the patterns of argumentation. Kitcher's model is the one that expresses in the clearest way the fact that, if we are to take seriously the profiled units of knowledge appealed to in an explanation, the quality that validates an explanation is to be delivered at a higher level than the information reflected in the explanatory mechanisms.

On the other side, this ambiguity in pointing to the right bearer of an explanation, as it is reflected by the above discussion, could be blamed on the too restricted ways of characterizing the knowledge into play, i.e., only by syntactic or semantic means. The limitation could be lifted by admitting a more pragmatic conception for the units of knowledge. Such proposals, as Kitcher's or Gardenfors' (in which the bearers are conceived as knowledge situations), explore this insight. But there is a price to pay here as the analysis of existing proposals revealed. The syntactic and semantic models of the structures of science tend to lose their importance as guiding means for the explanation analysis. It appears as though we are caught in a dilemma: either we use a detailed syntactic or semantic account on the structures of science but we fail to account for the pragmatics of explanation or we accept a pragmatization of the units of knowledge involved in an explanation losing this way the clear-cut structures of science as means for our investigation.

A way that could get around the dilemma will be, in my opinion, to take models as serious candidates for the bearers of explanations. They constitute important units of scientific knowledge and

[28] As it is developed in his book *Discovery and Explanation in Biology and Medicine*, and in other places.

an investigation in their ways of organizing scientific information and functioning in the process of scientific enterprise could provide the main guides in the search for a solution to the explanation problem. Given the earlier discussions, it is evident that we cannot limit our view to a strict syntactic or semantic conception of models. In fact, one of the main reasons to point to models as bearers was the fact that they exhibit pragmatic build-in components in their constitution. The fact that there is no general accepted account on models which recuperates their pragmatics, makes us to take into account seriously the possibility that we are not necessary going to search for a general model of explanation. The objectives we can afford for the beginning involve, besides the argumentation for the pertinence of getting the two subjects in closer contact, the search for the characteristics of the models that promote explanatory virtues.

First steps towards making the case for models – overcoming the prejudices

Perhaps the first necessary move that has to be done in the direction of the restoration of the respectability of models as epistemic and explanatory bearers is to distance us from the prejudices that overwhelmed this subject in the past. So, the first entrance in a kind of argumentation map would be engage some reactive arguments, which are to be directed against the depreciative view on models and its consequences as reflected in the inhibition of pursuing any inquiry into the explanatory virtues of the models. Of course, an argumentation map should not be reduced only to reactive arguments, even positive ones, which pledge for advantages in the search for solutions offered by the advocated perspective; and they should make out the main part of the map.

The reactive argumentation list is to be drafted following the main negative conclusions of the depreciative view. A first important item on the reactive argumentation list should be an argumentation against the limitations of models as revealed by the depreciative view, and considered to be responsible for the improperness of models as explanation-bearers. As already mentioned Hempel saw these limitations as the main hindrances in taking seriously the explanation provided by models in both cases: the analogical models and the theoretical one.

One may have doubts as to the necessity of such an argumentation once the received view on science has already become obsolete. The main claims of this conception came under strong attack from several sides during the last decades, so that almost none of them reveal sufficient pertinence for an actual project in the philosophy of science. Even though the major points of the conception were invalidated, the legacy of the view is still present in different topics. This is the case of models and modeling topic. Although alternative positions regarding this topic were voiced[29] quite from the beginning of the decline of the received view, they were still left as items of secondary importance on the working agenda. This fact contributed to the preservation of the ways in which these subjects were

[29] Mary Hesse and Marx Wartofsky are the boldest authors in the ’60s and ’70s addressing the modeling subject from different views as the received view.

conceived to be treated, and the place they occupy among the topics in the philosophy of science. Furthermore, even though there are no more major battles to be carried out on such subjects, a cleaning of the way from remaining prejudices will enhance a better understanding of the newly opened perspectives.

Therefore coming back to the problem of the limitations of models, we could differentiate several additional kinds. And although they are intimately linked, and a general argumentation against them could be relevant for more than one simultaneously, it is worth to spell them out one by one. Thus, a first type of limitation is a temporal one. This limitation is brought to light by the paradigmatic example invoked when dealing with the subjects, i.e., the example of the mechanical models for electromagnetic phenomena as they were developed in the 19^{th} century. It finds its expression in the claim that models are only temporary limited means of scientific investigation of which the scientist would dispense with when it arrives at a full theoretical representation, such as a law or a full theory for the phenomena under study. This aspect of models appears to be strongly manifested when taking into consideration especially the heuristic use of models. It appears also to find a more natural expression through the analogical models rather than through the theoretical ones, which seem to enjoy to some degree the status of a theory.

A short parenthesis: in as far as the semantic conception is concerned we encounter a dissolution of the problem, since the semantic concept of model abstracts from their temporal dimension. This will favor a better solution on explanation than the traditional one in case of an explanatory bearer without a temporal dimension.

Now, I think we can pertinently ask in what sense does the constitutive temporality of the modeling constructs affect the attribution of an explanatory power to models? There is no explicit requirement in the traditional program of the received view to ask explicitly for bearers of explanation that abstract from temporal aspects. Rather some imposition of this kind is implicitly contained in the requirements for the explanatory schema and it can be read out indirectly. In this program, the explanatory virtue is warranted by the status of the bearers, i.e., of the laws of nature, which have to be cited in the explanation. The nomic necessity is the core feature of this status, which validates the explanatory schema. This necessity implies the above atemporal character of the theoretical constructs. The critique on the traditional program pointed exactly to the difficulty of justifying this status; this way a solution for the explanation problem was made dependent on a satisfactory answer to a deeper philosophical problem, which involves some metaphysical commitments. The pragmatic turn in the explanation-debate could be seen as an indicator of the departure from such an interpretation of the problem.

Besides, if we are to take seriously the lessons from the historical orientation in the philosophy of science (in fact the participants in the later decades of the debate did), we have to make place for a temporal relativization in our accounts on explanation. The best example here is the one provided by the direct successor of the Hempelian program, Kitcher's account, which takes much care to articulate the temporal relativization of the units of knowledge in his model.

The temporal limitation of the models is a more general characteristic, which discredited the models from having a place in the logico-systematic account on science. Hempel did not even mention this kind of limitation as a hindrance for the attribution of explanatory powers to models. If we are to seek a reason for this situation I think it does not lie in the fact that the limitation does not constitute an impediment; rather it is caught and reflected under other limitations, which are more properly discussed in the register of the received view. The limitations Hempel explicitly considers are the ones of "limitation in scope" and "the approximate validity within that scope". I'll discuss them in the next paragraphs.

Limitation in scope or range of application has two sources according to Hempel: the idealization and the oversimplification inherent in the theoretical models. A direct consequence that can be drawn here is that laws as bearers of explanations are not subjected to idealizations or simplifications, since the last one will invalidate them for their explanatory claims. Such a position is not easy to defend, if we take into consideration the analysis of idealization[30] and its involvement in the constitution of the theoretical constructs (laws among them).[31] Or, if Hempel assumed tacitly that the laws are subject to these processes, but in a much lesser degree than the theoretical models, we need more reasons to understand why laws are explanatory and models are not. I would not pursue here an argumentation as to the fundamental role of the two processes in the constitution of laws, as I am going to touch the matter indirectly in several places, when presenting other positions[32] compatible with my view on laws and models. I want only to point to the fact that the way Hempel conceives idealization and simplification and their role in science is a naïve and limited. This conclusion is much bolder especially as far as idealization is concerned, since it appears that it is intimately involved in the structure of the laws. On the other side, the building of a model through the process of simplification and idealization is only a particular form of model building (as we'll see further). A closer look at the two processes, which prove to be so central in the theoretical business of science, leads us to the conclusion that they cannot be invoked in such a direct way as in Hempel's argumentation against the explanatory virtues of the theoretical models.

In discussing the theoretical models, Hempel refers mainly to the mathematical models in social sciences, even though he mentions at the end of the section that such cases are to be found in physical sciences too. One main reason for concentrating his analysis on models from the point of view of the social sciences could be the abundance of these models in fields that lack general formalized theories. Another reason could be a didactical one: these examples render in a more explicit way the nature of these kinds of models as containing simplifying assumptions of the abstract principles of theory. Some critical reading could also suspect that Hempel wants to circumscribe this type of models to mature sciences (in spite of claiming their manifestation in the physical sciences too) with the intention of keeping the domain of the mature sciences free of problems. A closer examination of the theoretical

30 See, for example, Laymon's analysis.

31 Cartwright's critique of laws as conceived by the logico-positivists is one of the most penetrant in stressing their inherent idealized status.

32 Cartwright's or Giere's positions.

models in the mature sciences would disclose the need of a more articulated view on models than Hempel's.

If we are to make explicit the modeling process in the frame of social sciences, the story is to be put down in the following manner (as it is met in Hempel's position): in trying to get some quantitative relations in the frame of a social theory that is usually caught in qualitative terms, we usually introduce idealized assumptions and simplifications. These assumptions make up a theoretical model. The image appears to be quite simplified and totally improper for physical sciences.

If we are to identify in physical sciences some analogue of this kind of model-construction, we could find a parallel in what Readhead[33] presents as a special type of theoretical model in physics, the 'impoverishment' model. Readhead's analysis on models presents the building of this kind of model by simplifying, both in a physical and mathematical sense, the equations governing the theory with the purpose of getting some more tractable equations, i.e., one with exact solutions. The model adopts also some idealizations that make it false in comparison with the true theory (an analogy seems to be justified in the case of, for example, the model of the rational agent in the economic theory). But Redhead describes also another type of theoretical model, the 'enrichment' model, as he calls it. Such a model appears in case of a theory, which specifies only general constraints, of symmetry for example (as it is the case of axiomatic field theory in physics), and it finds itself in need for more structure in order to be applied to a particular field. The model provides the necessary details and in this sense it enriches rather than impoverish the theory's claims. The assumptions of this model cannot be seen as simplifications of the theory's one, but rather as specification and addition of more content needed for the application of the theory. Such a procedure runs intuitively contrary to the one in the simplification case.

So, what other conclusions could we draw from above? Of course, it does not follow that the 'enrichment' model has a bigger explanatory power than the theory (though the first is an enriched version of the second), since the latter's explanatory power is claimed to be also directly determined by the extent of the range of application. But we can conclude that the way Hempel conceives the theoretical model generated by simplification and idealization is a very particular one. Furthermore, we can state that the conclusions drawn on this particular type of model proved not to be unproblematic if extended to other types of theoretical models. We can "restrict" the application range of the theory (if we are to hold to Hempel's proposal), by enriching the theory's assumptions in building a theoretical model. So, there is no such straight relation between simplifying the axioms of the theory and reduce its range of application. More probable, the larger range of application of the theory is rather due to an unspecificity, a "scarcity" (in Cartwright's words) in the determination of the exact reference class of applications.

If spelled out more carefully, the concepts of idealization and simplification not only have bad consequences for the articulation of Hempel's argument, but also for the notion of range of application.

[33] "Models in Physics", in *British Journal of Philosophy of Science* 31 (1980), pp. 145-163.

Now, if we are to take seriously Hempel's claim and try to make it more plausible, we could ask in what sense could we say that a theory has a larger range of application than a model?

For the received view on theories this claim is straight, since the range of application for a theory is to be rendered by the range of application of its laws, and this means that it is in a sense unrestricted. The range of application of the models is, under this view, confined to the particular systems they are designed to account for. These systems exhibit particular instantiations of the lawful relations handled by the theory.

The situation is not the same from a semantic point of view on theories. One direct possibility of identifying these ranges is a cumulative one. Since theories are sets of models, their application range will extend over the ranges of all its models cumulating them. If we hold on the identity of the laws and theory's application ranges, we can make sense of Hempel's claim. But there is also another way of looking to this problem, if we granted a more detailed view on the theoretical constructs (by separating different levels of theorizing which were conflated in the above simplified views).

The ranges of application of laws and theoretical models are two different things not due to a difference expressed in a quantitative way; but because they pick out different entities in the scientific enterprise. I'll make recourse to some elements of the semantic conception of theories in order to get a sharper clarification of the difference I'm pointing to.

According to this conception, the objects in the models are abstract, having clear-cut proprieties (those attributed by the theory), and satisfying the relations expressed in the theory's claims. The laws of a theory, having the form of statements in the theory, describe therefore the behavior of the objects in the models. The range of application of the laws is made out of the domain of these objects.

On the other hand, the empirical systems (and their objects) are not subject directly to these laws; rather they are subject to the theory's laws plus other stated or unstated provisos, i.e., statements, which are more or less expressed in explicit theories. They are rather brought together[34] in the entity of a model.

The point can also be made by using the difference Ronald Giere makes between two different processes of assigning reference to the mathematical symbols of the theory. One is the process of interpretation, as Giere calls it, the linking of a mathematical symbol with a general term. The other is the process of identification, which means the linking of the mathematical symbol to a feature of a specific object. Using Giere's example we *interpret* x from the formula of the force for a linear oscillator $F = -kx$ as the displacement from the equilibrium position; meanwhile we *identify* some features, i.e., the distance of a real specific oscillating system with x. The two functions would define two different notions of application with different ranges.

Let us turn briefly to the other direct statement against the explanatory pretensions of models, the one in which Hempel points to the "approximate validity" of the explanations delivered by models. This can be directly understood I think, as pointing to the failure of an explanation through models to satisfy the requirement for truth of the premises of the standard deductive nomological explanatory

[34] Boumans made this point forcefully in a recent paper which I'll discuss in the next chapters.

schema. I'm tempted to recall here Cartwright's critique directed against this requirement. The critique seems to me pertinent in showing that we can dispense with the truth. In consequence, approximate validity would not constitute a drawback in attributing explanatory virtues to models.

Concluding the above discussion it could be said that it is understandable why Hempel was so reluctant of involving models in his analysis of explanation. First and above all he was working under the received view in philosophy of science, view that conceived models as totally irrelevant for the logico-systematic investigation of the nature of science. But taking into account the post logical empiricism developments in the field, we could claim that there aren't any good reasons, less some knock-out arguments, to reject as inappropriate an approach on explanation topic by involving the modeling topic. I'll try in the further section to unfold this approach by investigating some possible positive ways of making the point.

Explanation in the pragmatic approach to models

In considering positive ways in making the point for the explanatory pretensions of the models, I'll discuss first of all a modality of doing it used by the pragmatic-instrumentalist conception on models as it is expressed in Morrison's approach. I'll try to see how pertinent the argumentation ways of this conception are in pleading for explanatory virtues of the models, since this position appears to make a very similar point to mine.

Margaret Morrison argues for an autonomous status of the models, which should be taken seriously when considering the scientific practice. Models act as autonomous agents in the production and manipulation of scientific knowledge. The autonomy is claimed in the first place from theories; but other constructs involving less theorization, as for example the phenomenological description of phenomena, are included.[35]

The autonomy she pleads for is reflected by many aspects of scientific activity; the explanatory aspect being one particular aspect among others. The importance of this aspect, on the other hand, is emphasized by the central role given to explanation in the argumentation. As she puts it,[36] by declaring it one of the core claims of her position: "that it is models rather than abstract theory that represent and explain the behavior of physical systems."

For M. Morrison the autonomy of models is the result of two different components:[37] first, their functioning in a way independent from theory, and secondly their construction with a minimal reliance on the high level theories. Actually, there is the second component, as she explicitly states, which is ultimately responsible for the first one, i.e., the independence in functioning. On the other hand, her

[35] Even if the point regarding the phenomenological descriptions is not explicitly made in this form, it can be seen as a consequence of her argument that the phenomenological models are not free from the influence of a theory.

[36] In *Models as Autonomous Agents*, p. 39.

[37] In *Models as Autonomous Agents*, p. 43.

argumentation makes appeal more to the first component of the autonomy. We can notice that this first component bears directly on her pragmatic-instrumentalist view on models, while the second component is in a way less dependent of this view. The explanatory function of the models is the most important concretization of the first component.

In order to illustrate how the autonomy is realized in various ways, M. Morrison provides us with many suggestive examples from different domains of the physical sciences: from the much discussed model of the mathematical pendulum to Prandl's model from hydrodynamics or the models for the atomic nucleus. Some opponent might object that much is left implicitly caught in the examples, for she does not succeed to extract and articulate in a detailed way some general features responsible for the explanatory virtues ranging over all these different kinds of models. This fits in a good sense the spirit of her pragmatic and historical oriented approach.

The grounds for the explanatory virtues of the models are given in a general characterization: "The reason that models are explanatory is that in representing these systems they exhibit certain kinds of structural dependencies. The model shows us how particular bits of the system are integrated and fit together in such a way that the system behavior can be explained." Someone taking a critical position could object that the reasons meant by Morrison to explicate the explanatory aspects of the models can be read out as disguised versions of the already well-known accounts of explanation.[38] Integrating and fitting different "bits of system" is what the core intuition of the unification account on explanation points to. Expressing some structural dependencies of the represented system can be seen as providing some law-like patterns of phenomenal order. Actually in the next paragraphs Morrison affirms the idea of explanatory gain by expressing law-like relations. So, the above-mentioned characterizations of explanation could be only particular instantiations of some explanatory schemas already well articulated in some explanation accounts.

But such critics could be accused of being on the wrong track since they ignore the main intention of the account. Morrison explicates further that: "a model explains the behavior of the system because it contextualizes the laws in a concrete way"; and further "the model makes the law-like relation evident in a way that the abstract theory cannot." It is therefore rather the particular way of instantiating of an abstract law-like relation that constitutes the core of an explanatory virtue of a model. But how are we to render more explicit this way of making evident the law-like relationship?

This specific "way" is to be rendered explicit by the scientific context in which the model was built and functions. The fact that this way cannot be articulated in a more general form is directly linked to her thesis that each model expresses a different type of representation. Since the explanatory role is a function of the representational features of the model, and according to Morrison we cannot isolate a sufficiently general form to characterize the representation that ranges over different types of models, there is no general way to express this function given by various representations. Ultimately, it means

[38] Some commentators as for example Tod Jones, regard Morrison's position on explanation as belonging to the ontic kind of approach, following Salmon typology; fitting events into natural regularities is, pace Salmon, a mark of the ontic conception on explanation.

that the limitation is due to the inability to distillate a general form for the representation provided by models; as Morrison puts it: "although the explanatory role is a function of the representational features of the model there is no one way to characterize the nature of that representation." We can notice here a reduction of the explanation problem to the representation one, which is further denied a general solution. We might interpret this as pointing to a local solution in the broader sense I'm pleading for.

Now, there are some pertinent critical points that can be made regarding Morrison's proposal. In general, it seems that a critical view will hinge mainly on the fact that a lot, especially in as far as explanation is concerned, is still left implicit or in need of more explicit articulation.[39] Of course, a meta-critical pertinent question will ask for the justification of further developments given the general pragmatics-centered frame in which the problem is settled; from Morrison's point of view we are certainly not in a position to pursue any further substantial explication.

Regarding the explanatory function of representation in the frame of the explanation problem, some major issues are still open. An account is needed as to how does the explanatory function of such representations in general arise, and how does it unfold in different kinds of representation. More specifically, given the already mentioned characterizations of the explanatory sources, an articulation is needed for whatever determines law-like pattern instantiation and unification to be caught in different types of representation, and for whatever makes out the difference between them. In fact, Margaret Morrison points to some distinctive features through which explanation is realized, but fails to provide us the modality they articulate the explanatory functions of representation.

As we may gather from the examples and the discussion on representation, she provides us with some suggestions of a sort of typology of representation. She distinguishes the "representation as mirroring" (the mathematical pendulum's case), or the mixed kind of phenomenological and theoretical representation as in Prandl's hydrodynamic model, or the ones exhibited in the models for the atomic nucleus; the last ones appear to be contradictory but nevertheless aim to represent the same system. These suggestions raise the legitimate requirement for some general criteria in identifying and comparing representations ranging over various models. Her claim that there is no one characterization for the representation has, according to this requirement, to be better specified.

Now the good news, I think, is tonic enough to trigger our engagement in a new perspective for the problem of explanation. There are some major qualities for harboring the search for an adequate solution. The main one is the fact that it provides us with a proper local perspective for such a project. An explanation comes into being under a certain context and draws its primary legitimation from the circumstances that justify the construction of the scientific modeling representation. Secondly, it has a kind of implicit dynamics, and therefore would require a suitable articulation of it. There are certain clues that point in this direction in my opinion. First, there is the fact that Morrison needs to embed her views in a functionalist modality of approaching the problems. She reflects this way the possible dynamics but, in my opinion, in an unsatisfactory way. The functionalist mode reduces the dynamic to a

[39] One of the common points of the critical reviews of her texts is the accusation of a lack of more articulation regarding the explanation subject.

single interpretation attenuating other possible articulations for it. Another clue is the stipulation of the model-building process as one of primary importance. It appears that she would regard as a viable point the reduction of the functionalist perspective to the building process. But she is not pursuing this agenda (besides an implicit reading which could be invoked from her examples), adopting instead the functionalist view as ultimate justificatory level. We are left this way with an open justificatory gap.

An indirect quality that could be read out even if it is only an unintended by-product of her account is a kind of "holistic" way of approaching the problem. First, it has to be noticed that it is only a particular kind, a limited one, which is certainly in need of further qualification. What I'm pointing to is the fact that she does not try to pick out a well-defined certain rationale delivered through a well expressed relation to constitute the core of an explanation; rather she regards the explanatory virtue as delivered by more aspects realized in the frame of the modeling construction. These different aspects contribute together to the realization of the explanatory virtue. I'll try in a later section to clarify further this component.

All these considered, I think Margaret Morrison undertakes some first steps in setting a pretty reasonable frame for pursuing a solution for the explanation problem in models. The fact that she does not make any more systematic attempts in these directions is caused by her working agenda, which does not seem to contain this entry. On the other side, her instrumentalist view makes out the main hindrance for a further development of the subject in the settled framework. By adopting a functionalist view as final one to such important topics as explanation or representation, her account rather obscures than promotes a satisfactory solution. Such a kind of pragmatic view as the instrumentalist one appears to be limited in its possibilities to promote satisfactory answers on questions as the above-mentioned ones. Given the richness of proposals in the explanation debate, the point for the explanatory autonomy cannot be made only by using the resources of a pragmatic-instrumentalist perspective; it has to engage some of the facts revealed by the very rich and diversified logico-philosophical analysis of explanation. Morrison appeals to such facts in her description of the explanatory features of models, but she fails to show how they articulate in rendering the explanatory autonomy of models. Therefore, her position is in need of such a pertinent articulation.

Now if we are to regard such a claim about explanatory autonomy of models from a general point of view, there are two modalities for making the case. The one is a kind of non-reductive way in which we could establish some explanatory virtues for the models but do not deny others on behalf of theories or other constructs. The second might claim this autonomy by placing the explanation only at the model level and deny it for other constructs. Morrison's conception belongs to the second sort. But the point is also directly linked with what conception on explanation one has or wants to adopt. My argumentation for the plausibility of viewing explanation as being connected with modeling brings me more in the second sort of strategy. Still I do not intend to exclude completely the first one; my main point being only that the coupling of the explanation subject with the modeling one could reveal more insightful resources in search for a solution to the explanation problem – this does not exclude the possibility of revealing some explanatory virtues at the level of other scientific constructs. Even so, I'll try to develop my argumentation as if there were no such other levels of explanation.

Now, to me it seems plausible that if we are to invoke a kind of thesis regarding the explanatory autonomy of the models from the perspective of the explanation topic in philosophy of science, it should be claimed against the constructs that were invoked in the main accounts of the explanation debate. So, the first kind of autonomy is the one from the theories. It is based on the fact that theories are empty in their explanatory claims without the appeal to models. So, it restates the thesis that we may get an explanation for a phenomenon when we are able to provide a model that represents it. This thesis manifestly formulated in its general lines can be found expressed by Cartwright. Aiming for other goals than an account on explanation, she puts the claim in a direct way in *How the Laws of Physics lie*: "to explain a phenomena is to find a model that fits it into the basic framework of the theory, and thus allows us to derive analogous for the messy and complicated phenomenological laws".

If we are to see how different conceptions on the nature of the scientific structures can accommodate the claim of models' explanatory autonomy, I think we can notice that it finds a natural expression in the frame of a semantic conception on theories. Since theories are families of models, it appears naturally to view the explanation of a phenomenon as fitting it under a model, which can offer a proper representation for it. This is a way adopted by Joseph Sneed in his paper *Structural explanation*, where he expressed the idea in the frame of the structuralist conception of science.

Nevertheless I think one may claim that the thesis could also accommodate quite well a syntactic view on theories with a more attentive view on models than the received view. Let's take as an example Achinstein's position on models, which are seen as sets of assumptions. His focus is primarily on theoretical models and to explain a phenomena under this scenario would mean to get the right set of assumptions – which makes out a model – for capturing the relevant features of the phenomena to be explained. So, it may be said that there is no restriction to a special conception of scientific theories that could foster the thesis of explanatory models.

Explanatory autonomy of models from laws is the other necessary part of argumentation. An adherent of explanation by laws would agree with the thesis that a model explains if it incorporates a law; he would point to the law as the main provider of the explanation. The idea could be extracted in an indirect way from Hempel's discussion of explanation by models; but the laws, which are incorporated into models, are usually some restricted kind of laws, not really the universal laws.

The reason for the explanatory autonomy of the models from laws finds its ground in the already mentioned distinction, which should be taken into consideration. This is the distinction pointing to the fact that laws represent actually the behavior of the abstract objects in the models not the one in reality. What the laws "explain"[40] is the behavior of some idealized objects in an idealized environment. There are lots of explicitly and implicitly accepted provisos that are used to settle the stories in the models. On the other hand, the explanatory resources for the phenomena are rather to be found in the models, which integrate in a direct way the representations of phenomena.

[40] In this case it is in fact an improper way of talking of explanation.

In the semantic view camp even if there is no strong claim made for explanation by laws, there are authors whose position will accommodate better the thesis of autonomy. Suppe's account on explanation or even Sneed's structural explanation could tolerate better the autonomy claim than for example Bartelborth's account.

Enhancing the plausibility of an inquiry into explanation through models

In this section I'll try to point to some facts, which are to speak for the plausibility of pursuing the analysis of explanation with an eye on models. These positive points should enhance the worthiness of pursuing the project of placing explanation at the level of models. They appear to be general guiding clues to be followed, so that the project could provide some answers for the problems raised by the explanation-topic. They constitute this way a kind of framework, in which further development should take place.

The first point I want to make concerns the material and formal conditioning on explanation and their mixture. We would expect that viewing scientific explanation as involving primarily some modeling issues will open a better perspective than any other inquiry strategies into explanation in reflecting both the formal and the material conditioning of explanation and especially how they combine. But before bringing the main reason for this expectation I'll try to make explicit what I understand by these conditionings.

The formal conditioning of an explanation concerns that type of determination that can be spelled out into formal patterns, i.e., the forms an explanation can take. Different accounts on explanation tried to pin down these forms ranging from a unique general one, as in the classical approaches on explanation, to domain-specific ones. The most plausible approach, given the present state of the topic of explanation appears to be the last mentioned (which in my opinion requires some significant enrichment with interrogative elements). But there is an important rest of conditioning left, which falls outside such forms. The material conditioning comprises that kind of determinations that resists being caught in forms. It founds its official expression in the explanation debate under the topic of the pragmatics of explanation. As the state of the debate on explanation in the last two decades revealed, this kind of conditioning plays a central role in determining the explanation. Even if the explanation problem was initially projected to be solved mostly under the first conditioning, the authors of the later decades called for featuring it mainly under the last conditioning; this fixed the agenda of solving it under a more pragmatics oriented perspective.

Models as scientific constructs exhibit particular characteristics in comparison with other scientific constructs – laws or theories. Models[41] exhibit a more heterogeneous nature than the other constructs by integrating ingredients of very different sorts. The important point for my argumentation

[41] As I'll discuss in another section, Boumans' account made this explicit.

now is the fact that the pragmatic and non-pragmatic factors are blended-in to form an efficient epistemic artifact.[42]

Regarding the formal aspects, models can include and integrate formal patterns belonging to different theories or ad-hoc developed ones that do not belong to any theory yet. The formal expressions taken from theories are subjected to all sorts of modifications: corrections, additions of supplementary terms or deletion of some existing terms. The selection and modification of the forms is determined by the pragmatics of the modeling processes. As noticed from the early reflection on science, the integration of these forms in the frame of the model does not happen under a strong 'logical coordination', as Duhem puts it. Aspects related to the pragmatics of scientific enterprise from different levels influence this kind of 'epistemic' coordination realized in a model. But, on the other hand, what is important for my point is that the forms integrated in the models are the proximate forms to be appealed directly for explanatory purposes and not the unspecified, more general one from the abstract theory.

As for the material aspects of the models, Knuutila and Voutilainen[43] in their characterization of models as epistemic artifacts describe them as "materialized inhabitants of the inter-subjective field of human activity". This kind of materiality is understood by the authors as pointing to the fact that models are "things that have their own construction and thus their own ways of functioning", restating in fact Morrison's position. Even if it appears that the way they clarify this materiality is a bit confusing, I think that the point made is worth to be stressed. So the main meaning intended under the notion of the materiality of the models is that in their construction and their uses models integrate and assimilate all sorts of elements falling in the pragmatics' register. The pragmatics is this way "built-in" as Boumans claims with regard to their justification. Models, it may be said, reflect the pragmatics at work in the scientific enterprise. They are therefore a better medium in which to consider the explanation problem. The major important gain to be expected from this strategy is one that should reduce the artificiality of adding the pragmatic dimension to the scientific constructs that were investigated initially in a pragmatics-free frame.

The choice of viewing the explanation connected with the modeling topic is also meant to avoid the extreme solutions that induce an implicit bias in the subject treatment. In this sense, explanation by subsumption of facts under laws, as in Hempel's model, as well as explanation as unification, as Kitcher's account pleads for, looses sight for the material conditioning of explanation stressing the formal one. On the other hand, explanation conceived as an answer to a question (usually why-question) in the way the interrogative approaches put it, looses sight for the formal determination of explanation for the sake of stressing the material one. A reasonable account should avoid such a biased solution. Explanation by involving models presents therefore better chances for avoiding such reductive views.

---

42 As Knuutila and Voutilainen name it.

43 In *A Parser as an Epistemic Artefact: A Material View on Models,* the presentation at PSA Conference 2002.

This is only a very general way of making a point; it is still in need of a more precise articulation. But on the other hand, I think this is an important clue for recourse to models in the explanation problem. A developed account on explanation, taking models as directly involved in clarifying it, should make use of the above mentioned advantage of models over theories or laws.

Another point could be the fact that models exhibit the right proportion between particularity and generality. Between the particularity of a phenomenon under investigation and the generality caught in the theory's principles, the models act as the mediators between theory and phenomena. As I already pointed out, M. Morrison pleaded for the view of conceiving models as mediators. The position came under attack that it did not provide enough strong criteria to identify the realms between which the models mediate. The request appears to me to be justified; but besides this critique the fact that models occupy an intermediate position between the generality of the theory and the particularity of the phenomena retains enough pertinence to justify the search for further articulation.

For this purpose they have to provide the necessary blend of the two aspects accommodating the generality of the theoretical claims to the specific phenomena under study or otherwise said to recuperate the relevant particularities of the phenomena for the scientific theorization. How does this speak for the placement of explanation at the model-level? The explanation problem is a philosophical problem that in its essence addresses the quest of the way our scientific constructs (which fall into the realm of generality) do recuperate the particularity of the real phenomena. Pointing at models as bearers of explanations focuses our attention on the constructs that are mostly in the business of blending the general with the particular.

A very important point to support my position could be made by considering the fact that important dynamic aspects of scientific inquiry become manifest at the level of models. One of my major dissatisfaction with the existing accounts on explanation is the static character to which they are subjected. This character enhances the view of explanation as an application. This character is especially expressed in the static way in which the knowledge is appealed in the process of searching for an explanation.

Models were already acknowledged in their quality of being involved in extending the theory into new domains of phenomena or in participating in the building of a new theory. Even the adherents of the depreciative view on models recognized this quality. Under a functionalist view, it was expressed through the heuristic function of models.

One could retort that this is only a specific function of some special type of model. In fact the logical positivists hold a very narrow view of models. One might expect that such a particularity could not be recuperated in a more general account on models as for example the semantic view.

French and da Costa[44] provided a solution which not only makes place for this feature in such a unitary account of modeling, but it makes also a good point for rendering some dynamic aspects of scientific practice. By developing the notion of partial structures, they built in the very nature of the models the possibility to account for the dynamicity of these constructs and further of the theories. So,

[44] In *Models, Theories, and Structures: Thirty Years On*.

this feature of models appears to be so central to them that it is recuperated across different approaches on scientific structures.

Another way of making the above point is the following. Based on the fact that models are subjected to lesser constraints than theories (Weinert[45] argues explicitly for this thesis), the first are also much easier subjected to be modified than theories. For Duhem the laws in a model are much looser coordinated than the ones in a theory; i.e., the last exhibits a strong logical coordination of the laws. Or putting it into the most recent way of reflecting the heterogeneous nature of models: different ingredients are baked into a model (Boumans). They are brought together not under a strong logical coordination but in a looser one. This can accommodate better than a strict logical coordination, the dynamics implicit in the scientific activity as expressed by the losses, changes or adding of new ingredients.

Inhibitions still left?

Even after these efforts to increase the plausibility to link the two topics, I think, there are some fears still left to be named. The fears could be expressed by someone still skeptical of the project that I'm pleading for. Identifying them does not bring necessary some new items into play above the ones already mentioned, but can throw a light on the subject and open new perspectives. In general, I think these fears can be seen as being triggered by the relaxation of the requirements set by the previous explanation accounts. The best expression of this is of course to be extracted from Hempel's model, but Kitcher spells them out in a systematic and more actualized way too.

One that comes first of all to mind is a kind of fear for an arbitrariness inherited by the models as scientific constructs, arbitrariness which would contaminate the explanatory process too. We could recognize it in Duhem's warning against the danger of the free proliferation of the models. This means that under more loosely constraints, as is the case in modeling processes, we could generate models in a relative arbitrary way.[46] It is relative because there are still constraints that are to be fulfilled when building models. Somehow, this arbitrariness will bear on explanations if we are to take models seriously in this business. Ultimately, it will undermine the very idea of the explanation topic (one could object in extreme), that there is something to be identified as *explicatum* for the scientific explanation, i.e., that there is an explanatory structure sufficiently general to be identified.

But, not all approaches on models 'exhibit' this fear. This kind of feature could be played down if the models are conceived to reflect strictly the theory as in case of the semantic conception of theories. The "free proliferation" is in this way kept under a stronger control. The principles of the theory are the ones that regulate the status and generation of models.

[45] In *Theories, Models and Constraints*.

[46] This arbitrariness is to be seen in opposition to the systematicity exhibit by theories.

The above reaction can be put in terms of the old way of talking by regarding the difference between possible explanations and real ones. These arbitrary generated models will eventually contain tentative or possible explanations from among which the real explanation can be chosen. To take the modeling framework as basic for an inquiry on explanation would mean to fail to identify the true problem of explanation – what makes an explanation a real one rather than a possible one. Of course, this way of putting the problem rests on the assumption that there is something sufficiently general to be searched for as determining what makes an explanation real.

I'm afraid that we have to accept the lesson from the pragmatics of explanation as argued by, for example, van Fraassen and renounce the pretension of searching for such necessary and sufficient criteria caught into a clear-cut (syntactic or semantic only) schematism. There is such a relevance relation between the questions, which asks for explanation, and the answer, i.e., the explanation, to be identified as the last essence of explanation. This relation varies contextually, as explicitly shown by van Fraassen, and resists any representation into general forms.

On the other side, regarding the loss in stringency of the conditions imposed on models, I think that there are sufficient constraints that regulate the models' status. As Weinert[47] (and other authors) suggested, the spectrum of constraints is quite large. It is possible that the explanatory virtue should require different accents in different situations, so to emphasize some constraints according to the situation at hand. So for example, the empirical requirement of adequacy with the data could make out the principal reason for a model to be explanatory but there is also possible that some other constraints (or a combination of them) should be required to be emphasized in order to make the model explanatory. But perhaps this is a too bold conjecture that has to be more carefully analyzed.

Another fear (and a way of formulating it) that seems to be important is the one that could still invoke a kind of insufficiency of models to provide explanations. In the previous paragraphs, I came close to it, but didn't name it explicitly; it is in fact closely linked (or it follows naturally) to the above mentioned topic. To put it in a direct way this fear would refer to the insufficiency of the models to provide enough "restrictions" to determine an explanation. It calls further for the need of an extra-ingredient delivered by some other means than the ones falling in the models domain. It is in a sense another way of expressing the restraint against models due to their limitations. Such examples of extra-ingredients are unification – which should be delivered by the theories or other bigger units of knowledge – or causal mechanisms, which do not require any specific form of scientific knowledge.

In principle I see this request for an extra-ingredient as too demanding and any account that promotes it as still being under the influence of the old ideal of the explanation topic. The demand for extra-ingredient points to some more stable, systematic features of science than the one expressed in the models. I cannot see any good reason for refusing such a feature to be instantiated at the level of models. The most important thing is not that this last mentioned fact happens, but that the frame in which it takes place – the modeling activity – offers the best perspective to express the openness and creativity involved in the scientific enterprise.

[47] As discussed in his paper *Theories, Models and Constraints*.

Not last in importance, is a fear generated by the variety and plurality exhibited by the modeling topic – usually expressed by the variety of the functions, which are attributed to them. It undermines again the uniqueness of an explanatory schema, ultimately playing down the pertinence of the project of the explanation subject.

Nobody can deny the important roles models can play in scientific enterprise. If explanation is one among others, it does not mean that it is less important. But of course it does not cast any more a long shadow as it did in the past, as one of the most important subjects of the philosophy of science.

From the previous paragraphs, one could read out that the fears we have discussed above are reduced ultimately to fears generated by renouncing the well-coined frame of investigating the explanation problem, i.e., the 'classical' settings of investigation. After the dismissal of the 'big stories' about science and with the ongoing threat of dissolving any logico-philosophical pertinent discourse, it is time to take seriously the rich but puzzling parts of the scientific enterprise; as by taking seriously the topic of scientific modeling we will engage in the investigation of such parts.

Chapter 3

Placing scientific understanding in the new frame of inquiry

The main effort of this chapter centers on the evaluation of the ways a central concept for the explanation problem – the understanding concept – could be accommodated to the new frame of investigation. The importance of the concept is even more emphasized by the fact that in some approaches it was seen as being almost equivalent with the explanation concept in the sense that the account on explanation is to render further an account on scientific explanation and vice-versa. This fact reflects a certain attitude towards the similar role played by the two concepts in the economy of scientific inquiry. It also justifies my focus in tackling first of all this concept in order to see what new valences it acquires, and what clues does it offer in situating us as to the explanation problem. Nevertheless, philosophers engaged directly in the explication of the explanation concept acknowledged a separation between the two notions.

In order to engage in the above effort, a number of other concepts usually deployed in the articulation of an account on scientific explanation will have to be reevaluated in the new light. The lines of the investigation follow the mentioned ideas that might implement better the suggested features of a plausible account – to be local and to capture the dynamic aspects of the explanatory situation.

***

It is to be noticed that few addressed the issue of understanding in a modeling context. Even the modelists, i.e., the adherents of a model-oriented approach, didn't give it a place in their reflections. Some of the reasons behind this fact are similar to the ones related to the explanation through models. But there is also the more general attitude towards the topic of understanding that has to be taken into consideration, as I pointed in the historical chapter, a refrain from addressing the theme of understanding in the philosophy of science in a more systematic way. The situation changed a little in the last one and a half-decade during which some explicit articulated proposals appeared on the stage. Thus one could notice that there was a growing interest in this last period towards a more careful consideration of the topic.

As it was emphasized in my earlier chapters, the relation between explanation and understanding was usually considered to be a very close one. Some authors even indulged in a synonymous use of the terms (of course, this also depends on the register of the inquiry). Nevertheless, the relation was scarcely made for a subject of explicit reflection in the literature. The accounts on explanation took it for granted in the sense that it regarded it as a primitive fact, and incorporated it into the approach without feeling the need for a justification. The relation was usually thought to bear not so much on the articulation of the account. This attitude can be also read out in the case of the unification approach that claims to offer an explication for both concepts and the connection between them.

There were authors pointing to the divergence of the concepts and to the need to account for it. In this case, one may impose the requirement that a complete theory of explanation should specify when and how the two concepts are identified and when they are decoupled. Lambert is one of the few among the philosophers to inquire explicitly on the necessity and sufficiency of the relation between the two, offering also an explication for it. His conclusions[48] are negative regarding both claims: scientific explanation is neither necessary, nor sufficient for yielding scientific understanding. The relation between the two is viewed under a certain construal, i.e., that a scientific explanation yields scientific understanding and it is analyzed in an interrogative frame. In this sense, an answer to a why-question that qualifies as scientific explanation yields scientific understanding if and only if it provides an answer to the question "how does the *explanandum*-fact fit into the theory?" His view was further developed (in a joint paper with Schurz) in a well-articulated model of scientific understanding. I'll evaluate the potential suggestions for my inquiry of this model later in the chapter.

Spohn provided an alternative answer to the question of necessity and sufficiency of the relation as a reply to Lambert's view. By drawing on a specific kind of explanation, the causal one, and developing it in a particular frame, i.e., a belief-revision type of approach, he renders the relation between the two as necessary and sufficient. The understanding concept is construed in a particular way as providing stable reasons. A first moral that comes out of this debate will point to the fact that we can construe the relation under different ways and accordingly recuperate necessity or sufficiency or both.

We could subsequently ask in what way a model-oriented view would reshape the above problem, by imposing some limitations, and constrain the ways of searching for solutions. On one hand, we can reasonably expect that, by imposing some constraints coming from types of models and modeling strategies. On the other hand, given the diverse spectrum of meanings covered by both, the explanation and the model concepts, the prospects for a general theory accounting for the entire variety look rather slim. Our hopes should be censured accordingly.

From an historical perspective on the topic I'll characterize the actual approach on understanding as an explanation-centered one. The subject of scientific understanding was addressed only in an indirect mode, through the inquiry into the explanation problem and in strong connection with it. It made therefore for a by-product of an account on explanation. Even Lambert's line, though he rightly points to the necessity of getting independent accounts on understanding and explanation, in its fully developed version (as expressed in his and Schurzs' model of scientific understanding) could be seen as being an explanation-centered one.

Acknowledging the above facts I'll rather endorse Humphreys' claim[49] that a much richer field of investigation opens by trying to approach understanding directly. This does not mean to abandon any reference to explanation or how understanding is related to it. Rather I see Humphrey's urge to be mainly an attempt to place understanding in the central position as *analysandum* of the inquiry rather than to get it as a secondary result. This contradicts directly Friedman's view that it is not reasonable to

[48] See *On whether an answer to a why question is an explanation if and only if it yields scientific.*

[49] See his *Analytic versus Synthetic Understanding.*

approach understanding in isolation. We can find what understanding is according to Friedman, only by finding out first what scientific explanation is, though he acknowledges the fact that general ideas on understanding give us guidance for our inquiry on explanation. But Friedman does not provide us with any reasons to establish why a direct approach should be unreasonable. In subtext, some could accuse the same *horror pragmatics* that motivated the rejection of understanding under the received view.

To resume briefly the main characteristics which constitute the way understanding was approached in the main accounts, we can mention the following points: the explanation-centeredness of the approaches, the lack of explicit reflection on the necessity or sufficiency of the relation between understanding and explanation, the existence of a "taken for granted" attitude towards the strong link between the two concepts and the global character of the register in which understanding was considered to belong.

Situating the explanation problem in a modeling view does not necessarily force the abandonment of the above characteristics – with a major exemption: the globality requirement. I'll discuss further in the chapter in a detailed way this last point. But I do think that abandoning some of them or making more flexible the attitude towards them is a move that has to be undertaken when approaching understanding and explanation in the modeling context. My hope is that the particularity of different modeling contexts will indicate which has to be retained and in what form.

The chances of the traditional accounts on explanation

In the next paragraphs I'll evaluate the already proposed concepts of scientific understanding as found in the classical accounts in the perspective of extending them in the register of models. I'll try to point to the main features that hinder more or less a successful extrapolation of the concepts to the modeling context, making implausible a further development on these lines.

I'll turn first to Hempel's characterization of understanding as nomic expectability. We might ask what do the terms in the syntagma suggest to be the needed qualifications that could determine the specificity of models which exhibit such a kind of understanding. The nomicity requirement links it strongly to the DN-model of explanation; but allowing for a more relaxed reading, it could be taken as a reference to the laws incorporated in a model. The second term points more to the predictive qualities of the models. Some would reject this narrow identification of expectability with prediction since the first concept was intended to refer to a subjective aspect of the inquirer. But our effort is also one of objectifying the nature of understanding and in this sense we do move clearly from the beginning against Hempel's intention.

There are many examples of models that can exhibit the above characteristics. There are many types of models built as direct consequences of some laws, usually constituting direct applications of the laws to specific systems. Some of them are used also to provide us with predictions. But these characteristics do not tell us much about the understanding they provide or are quite useless in imposing some constraints to characterize it. So there is a large class of predictive models built in this way that are totally devoid of any intention to offer understanding; their main purpose is to provide a mechanism of generating possible data, playing the role of predictive devices.

Regarding the above last point, one could say that we beg the question since the critique of the DN-model already exposed the inadequacy of incorporating prediction in an account of explanation. So, if we would like to reveal the senses of scientific understanding and scientific explanation as caught in the models according to the DN-account, we should try to avoid the weak aspects. We could try to impose only the nomicity – i.e., the presence of laws in the models – as the necessary requirement for capturing at least a sort of scientific understanding provided by models and letting "expectability" aside as a subjective aspect, for example. This is a too weak constraint that has to be supplemented with some other requirements. But it will also force us to get more in the direction of scientific understanding through laws that, I think, will deviate us from the primary intention of this inquiry.

Friedman's and Kitcher's concepts

Friedman's and Kitcher's concepts of understanding could provide a richer ground than Hempel's suggestion, since their accounts on explanation intend explicitly to recuperate scientific understanding. We could expect that the main problem of extrapolating the concepts from these accounts to the modeling context lies in the way they are explicitly developed to incorporate a global dimension of scientific enterprise. For Friedman scientific understanding is "a global affair"; it does not address the intelligibility of an individual phenomenon, and has to do with the increase of the overall understanding of the world. This induces from the beginning a major difficulty in naturalizing it in the model frame. Models are scientific constructs sensible to the individuality of a phenomenon and run in this sense, in opposite direction than Friedman's sort of approach. The possibility of a pertinent transfer of some suggestions for inquiring on understanding in a modeling context will depend on how much the global aspect of the articulation can be attenuated. I will consider the issue of globality and locality in a greater detail later. Here I'll be concerned rather with the possibility of taking the main idea of the account as guiding in our inquiry.

Friedman will most probably reject any such attempt to address the issue of understanding through models since it openly violates the two requirements he sets at the beginning for the approach: to be general and to be objective. The specific nature of models would inhibit, according to the received view, any plausible attempt to satisfy these requirements for an account of explanation and understanding through models.

Scientific explanation and understanding are provided in Friedman's account via the reduction of the particular laws or law like generalizations to a more general law. This can happen quite well, as it is in fact the case in the frame of a model. More phenomenological generalizations or lower level law-like generalizations could be unified through reduction à la Friedman, to a more general law that constitutes the core of the model. There is no strong reason that could speak against such a reading; the constraints that Friedman assumes on the scientific entities – that are, broadly speaking, the ones from the received view – do not exclude it. One could rather react by invoking the fact that the kind of derivability involved in a model makes too much recourse to particular assumptions, and spoils in this sense the 'generality' of the proposal. This reaction is not really effective since many episodes of scientific reduction proved that such particular assumptions are often involved in the reduction between laws too. Among the models that contain such reduction will be some high level theoretical models. They could also be built for specific reasons – as for example heuristic ones – and some will make for good candidates for the status of a theory. The understanding they provide can therefore be similar to one provided through theories; a more careful consideration of this claim should be undertaken in the frame of theoretical modeling.

We would further ask what could then hinder us from seeing such an approach as being adequate for tackling the issue of understanding through models. A first limitation one could suspect is that the construal is too restricted to a special kind of model, i.e., one that contains such reductive relations between laws, in order to be applied to other types. But this objection can be played down under the proviso that we are not after an exhaustive account for all types of models. Nevertheless, even in the case of models that exhibit such a reductivist episode, the sense of scientific understanding provided by the model does not necessarily coincide with the one conveyed by Friedman's approach. The understanding of how specific laws are deductively obtained from other laws can be of a secondary importance, while the main sense of understanding provided by the model can quite well point in other direction. The main purpose of the model could be a heuristic one or a predictive one or of another sort. Accordingly the main sense of understanding it provides in the economy of our inquiry will not be given through the reduction episode it incorporates.

In Kitcher's approach the appeal to the global character of explanation and understanding impregnates even stronger than in Friedman's case the articulation of an account on understanding by models. I'll further evaluate some possible modalities of tackling the topic and expose better the reasons behind the above claim.

Another move will be to try to preserve the account and make room for models. The important units of knowledge that Kitcher's account considers are entire corpuses of knowledge as accepted by a scientific community at a historical moment. These are conceived as being systematized by sets of argumentative patterns. Other units of scientific knowledge, as theories or models, are constituted of instantiations of such argumentative patterns. In the spirit of the account, models could be claimed to exhibit explanatory qualities by virtue of incorporating some of the explanatory patterns.

The strategy would run counter the original purpose of our inquiry since it will strip models of their explanatory autonomy. Their capacity of providing understanding and explanation will be delivered top-down through the patterns that belong to the E(K) – the explanatory store over the corpus K. But even putting aside this fear in the hope we could save at least partially this autonomy, it looks to me that following this line will also open difficult problems without having a sufficient resourceful frame to handle them. One of the problems will be to account for the fact that models can incorporate patterns of various sorts,[50] not like theories whose identity are given by a specific pattern or set of patterns. A bold question that arises, would ask by virtue of which pattern is the model explanatory; or how could the explanatory quality of the model be built up as a whole from the different explanatory patterns it incorporates. Or how are we to account for the inclusion of patterns from non-explanatory systematizations, which contribute to the explanatory qualities of models.

Another serious problem will be the distortion induced to the nature of models (as well as of the theories) rendered only through a syntactic view; as we would be constrained by the sentential nature of the corpuses of knowledge as stated by Kitcher. Some people could try to argue that the non-sentential elements are not important for the issue of explanation. This sounds like a 'syntactic chauvinism' that brings back a strongly rejected agenda. Some enrichment of the frame idea could be tried by making recourse to particular stipulations incorporated in components of the patterns: the schematic argument, the filling instructions or the classification in order to capture better the specificity of some type of models. But I think we are confronted here with quite an unmanageable task, and I'm very skeptical that this move could be worked out into a reasonable solution.

We may try a more direct application of Kitcher's ideas to modeling units in the sense of taking them as referential units of knowledge for explanation instead of the corpuses. We would check consequently, how well the mechanism on which explanatory quality is built is transferable to the models' case. Of course, some of the difficulties exposed earlier related to the identity show up again, but let's bracket them for the sake of the argument. I'll take them as having a sentential representation and an identity that can be pinned down in this way and focus on the issues related to the mechanism. One of the major questions that arise concerns the relevance of the procedure for models. The criterion for comparison between the corpuses is the unification power of the sets of argument patterns, which does not seem to be the case for deciding on taking a model as explanatory. Other considerations than the unification power are also taken into account in the last case, and the unification power does not play a unique or essential role. A major reason lies in the fact that models are purpose-driven scientific constructs and therefore comparing them according to their unification power does not have an overall priority in evaluating them. How good a model accomplishes its purpose does not draw so heavily on how well its patterns unify – unless, of course, this is its main purpose (one may add).

Salmon's concept of understanding

If we are to consider further how the concepts of understanding proposed by Salmon will fit the bill we wanted, we will tend to disregard the first sort of them due to its similarity to the above discussed ones. This first kind refers to understanding achieved by showing how a phenomenon fits into

[50] In his 1981 paper on explanation Kitcher is aware of a similar problem regarding the use of patterns which do not belong to the E(K) but he does not address it.

‘the general scheme of things’, that is, into a ‘scientific world-picture’. Salmon follows in a more informal way the core intuition expressed by the unificationist approach, but unless the previous accounts it is a too loose[51] one. The previous objections related to the unification approach apply roughly in this case too. Models could be considered to be the vehicles for achieving this fitting, but in the absence of any explication of how to articulate the fitting-process, we cannot make much out of the suggestion.

The other type of understanding that Salmon points to, is achieved by revealing the causal mechanism at work in a particular situation. It is often realized through a model and it is central to some scientific domains. Later in the chapter I’ll address this topic in a more detailed analysis. Nevertheless, Salmon’s description proves to be too informal and is lacking in precise articulation of why or how we should build understanding in this case out of the knowledge of a causal mechanism.

Salmon provides[52] us also with a schema of different types of understanding among which the two scientific sorts discussed above are being classified. Since it doesn’t seem to be any place in the schema for a model-based brand of understanding – except eventually the causal type, we could ask how somebody would see it in the light of this classification. To deny completely the existence of such an understanding is to cut too harshly into the evidence of the reality. Probably Salmon will regard such models as vehicles that contribute to the building of the forms of understanding described. So they could provide a kind of pre-understanding and a partial one, also a tentative one, which should lead or enhance the classified forms. On the other side, building and using models engages all forms of understanding exposed in the schema, and we can view the understanding through models as a kind of mixture of the pure brands revealed. But to engage in a further clarification along such lines, on the basis of Salmon’s scheme, is an awkward effort and will make for an open agenda with a debatable plausibility.

.

Lambert-Schurz account on scientific understanding

The Lambert-Schurz’ account on scientific understanding is a unification-type of approach exhibiting quite an elaborate articulation and a broader intended range of application than the other two approaches discussed. The main lines of some of the critiques exposed previously remain valid, although they have to be properly qualified. I’ll evaluate further the account and try to reveal its more or less promising aspects for adopting it for my inquiry. In general, we will expect that the chances to adapt it proper to an account for a model-based understanding would be better than for the previous unification approaches.

The authors’ explicit intention is to accommodate different sorts of knowledge corpuses. They characterize them as cognitive corpuses, taken under a certain kind of representation – through statements – and which are also relativized to an epistemic agent AG. AG could be a person, a scientific

[51] Also pointed out by Trout in his critique of the concepts of understanding.

[52] See his text *The Importance of Scientific Understanding*, in *Causality and Explanation*.

community, or even a computer. Unless in the previous accounts, we could recognize from the beginning that some intended sense of the cognitive corpuses could recuperate the meaning of certain types of models, as it is the case of the corpus of information from a computer. I'll focus on this last case in more detail later in the section. For the purpose of keeping the section short, I'll not go into all the details of the account, and mention only the needed ones for the argumentation at hand.

To recall shortly the main lines of Schurz-Lambert account, understanding is conceived as a ternary relation between an answer A (to the understanding-seeking how question: "How does P fit into C?"), a phenomenon P (a description of a phenomenon) and a cognitive corpus C, relation that can be expressed as A contributes understanding of P relative to C. The answer A shows actually how P fits into C, which will change into C* after the addition of A. C is conceived as a pair of two sets: one of phenomena (K) and one of inferences (I), and the fitting of P is achieved means an argument ibs (ibs stands for 'in the broad sense') connecting P to some parts of K such that K's unification increases. Fitting into is described as connecting plus unification. The authors proceed further to articulate rigorously the above ideas. After defining the notions related to the answers, to understanding-seeking questions, they address the explication of the unification analyzing it as a comparative concept involving an ordering relation over the set of cognitive corpuses. To account for the contribution of different Ps to K, they attach weights to phenomena together with costs for different moves in the corpuses (and provide also an algebra of shifts and u-values). The full formal articulation is done with reference to a formal language (comprising at least first order logic) and a cognitive agent. Different types of connections are defined as well as different types of assimilations and dissimilations of Ps into the Ks.

As a consequence of their theory, the authors present us with a typology of understanding for the two distinct families of understandings: understanding-why and understanding-about, followed by a reflection on the relation between explanation and understanding conceived classically as explanations constituting the means for understanding. They show how we can select different views on this relation on the basis of their theory. The kinds of understanding, the two families: understanding-why and understanding-about and the consequences regarding the explanation-understanding relation are directly read out from the different ways of the "fitting into" description. We could draw on this typology in a modeling context if the more general assumptions of the approach would hold in a sufficiently pertinent way. I'll mainly address this point further in my discussion.

Basically, similar objections raised earlier apply here too. One is the limitation due to the statement view adopted. The cognitive representations in the corpuses are sets of statements mirroring the phenomena. No room is left for non-statement representations and this fact bears on the adequacy of translation to models, as it is the case in the previous discussed accounts. For example, the case of fitting a diagram or a graphic representation into a model will be viewed in the account as fitting a set of statements that translate the representation into the corpus. Or to take as example an adjustment of the size of a scale model that will enhance this way our understanding of the phenomenon under investigation, will make for a fitting under the discussed account, if the situation is translatable into a statement frame.

Another inadequacy is given through a sort of requirement of 'maximality of the information' (to call it this way) that is in fact an implicit assumption of the unification approach in general. It refers to the fact that a corpus should contain all the statements rationally believed by the agent. Or in other words, the cognitive corpus has to exhaust the agent's knowledge. This makes the interpretation in terms of models harder, since models are entities that are defined in contrast to some larger units; they do embed a sense of delimitation that is denied when we identify them with the corpuses. On the other side, to conceive them as fragments of these corpuses raises a similar problem to the one previously revealed. We could eventually try to extend the account to models by introducing a new kind of parameter that will induce the 'delimitation' of the models as units of knowledge with respect to the entire cognitive corpuses. But I don't expect this to be a move that will improve substantially the chances of the approach to capture the features of scientific understanding at model level.

The neglect of the pragmatic aspects of the models is probably a more general way of putting the objections. First of all, models as purpose driven units do harbor a sense of understanding as influenced by this characteristic. It makes for a difficult task, I think, to see how this could be articulated in an adequate way. Under a more generous interpretation, Lambert and Schurz do actually provide some modality to capture the pragmatic aspects. They are captured indirectly in the weights attached to the phenomena. But it is an implicit way of putting the things in the context of an inquiry into modeling issues and it fails to do justice to the importance pragmatics has in this problem setting.

Turning to the specific case of the computer-agent, it strikes us as a weird claim to attribute understanding to an episode of information manipulation through computers. What seems to me striking is the fact that this example runs counter a natural intuition regarding understanding – that it is to be attributed only to human subjects. The human agent is the agent that gains understanding by information manipulation via the computer, and this brings back pragmatics. Simulating or predicting the behavior of a system through a computer-implemented model provides the inquirer with understanding by interpreting it in the context of the human-driven scientific inquiry.

We could eventually take the computer implementation as a representation for certain purposes of a scientific model which is more comprising, but not to be reduced to the model in the computer. In fact it is not clear even if implementations are to be taken as being models, as the recent conference on *Models and Simulations* (Paris, June 2006) showed.[53] The implementation of a theoretical representation in the computer contributes to the exploration of the model by providing further scientific understanding. It is used for such purposes as exploration of the predictive power, or the discovery of new consequences towards which the model assumptions lead. The implementation enhances our understanding of the model that of course will further bear on the understanding of the

[53] John Simpson in his contribution to the Paris Conference 2006, *Models and Simulations*, argued that simulations should not be properly taken as models.

natural system. But the targets of the two understanding processes seem to be different and we have to be cautious not to lump them together.

To return to some general observations regarding the basic idea of the account: the cognitive corpuses may be seen from one perspective to be more general, more comprising than the knowledge caught in the models, and from another perspective too restricted as they are represented as sets of linguistic entities only. Therefore there is an inadequacy of applying the account to models from several directions. Nevertheless, as the basic idea of a phenomenon is being assimilated into a corpus of knowledge, it seems to provide an intuitive appeal, and in some cases can even make for a description as to how we gain understanding. We could quite well conceive the inference of a phenomenon by using a model, i.e., by drawing on premises from the model, as assimilation in Lambert-Schurz sense. But even in this situation, we can make the case that this will not always be the sense of understanding we are searching for. We do understand the assimilated phenomena, but it is possible that we might not enrich our understanding the model gives us about such kinds of target systems.

The sources of the difficulties revealed in the above discussion on the classical accounts on understanding are not to be reduced exclusively to the fact that they are tailored to capture the global features of the theories or, more general, of the corpuses of scientific knowledge. There is also a more hidden assumption at work – they are still tributary to a received agenda of dealing with such problems. This agenda is characterized by a *horror pragmatics* and fosters a more formal and unique solution to the problem. These requirements are to be relaxed at least in a prospective phase of the search for a solution in the context of models.

The inadequacy of the classical concepts of understanding as analyzed above calls for a reorientation in the ways of approaching the subject of understanding, especially if we want to naturalize it in a modeling context. Furthermore the move is desirable if we intend to approach it in a direct way. In this sense, it seems plausible and it can be helpful to consider a check for more generous attitudes under which we can situate the concept of understanding. I'll point to some suggestions in this sense and try to evaluate them, and what kind of promising perspectives they open up. The recent conference on the topic of scientific understanding (Amsterdam august 2005) revealed the main landmarks of the landscape. (P. Lipton's lecture delivered a good overview of these landmarks). The more generous perspectives are those that are not limited to a certain kind of general agenda, as in the case of the previous accounts. They are not "pragmatics-shy", nor even make heavy reference to it. Further on I will sketch and evaluate briefly a proposal of a pragmatic account and the reaction against it.

De Regt-Dieks account on scientific understanding – a direct and pragmatic approach

Recently de Regt and Dieks provided an account that may be viewed as a direct approach to scientific understanding. The authors' intention is to prove at the same time the unifying role which understanding can play for the particular explanatory forms revealed in the explanation debate – a goal that Lambert and Schurz were after in their account too. The paper develops also a critique of what the authors call the ‘universalist’ accounts on the concept of understanding by which they intend to characterize a type of account “that assigns a privileged status to one particular form of explanation by claiming that if an explanation of that form is possible than this is always the best explanation.”

I’ll present briefly the main ideas of this account, expose its strong and weak points and draw some first evaluation for the understanding through modeling. I’ll also try to draw some morals or pertinent points that can be taken further into consideration for an inquiry in the frame of a modeling view.

The goal of the account is not only to provide a certain explication of the concept of scientific understanding but to situate it also in the perspective of scientific enterprise, and plead further for its important role. In this sense the authors argue from the beginning for the claim that understanding is an important aim of science, contrary to other positions (e.g., Laudan) that deny this fact. To achieve understanding is seen as a general aim at the macro-level of science, i.e., science as a whole, meanwhile variation with respect to the means of achieving it appears at the meso-level – the one of scientific communities, or micro-level, the one of the individual scientists. The authors try to avoid this way the critique raised by invoking the historical variations of scientific aims. They also plead for the epistemic value of understanding by rejecting such positions as Hempel’s or van Fraassen’s on the irrelevance of scientific understanding for a logical-philosophical investigation of scientific knowledge. Descriptions and predictions are important aims emphasized by positivists, but science is also searching for understanding, as de Regt and Dieks stresses. They agree with the positivists’ that understanding is essentially a pragmatic notion, but this does not invalidate its relevance for the philosophical analysis of science. Their account puts pragmatics at the heart of understanding and it intends to show in this way that though it incorporates a subjective dimension, it can be accounted for in an articulated way.

A second main assumption lies at the bottom of their account, an assumption which states that scientific understanding requires theories and it is explicitly caught in the first general criteria they provide. According to it – called the criteria for understanding a phenomenon (CUP) – a phenomenon can be understood if a theory for it exists, that is intelligible. The second criterion is meant to explicate the qualification of intelligibility <u>of</u> a theory. According to CIT (the criteria of intelligibility of a theory) “a scientific theory is intelligible for scientist S (in context C) if S can recognize qualitatively characteristic consequences of T without performing exact calculations.” In case of non-mathematical qualitative theories “exact calculation” is to be replaced with “complete logical argumentation”. The authors exemplify their ideas by such scientific episodes as Boltzman’s qualitative presentation of the kinetic-molecular nature of the gases or the potential-vortices type of thinking from meteorology, used in describing the behavior of cyclones. According to the two authors, such a qualitative analysis provides an understanding of the phenomena “before embarking on detailed calculations.”

De Regt and Dieks make room in their account for the plurality of means that CIT can accommodate, through which understanding could be achieved. The qualitative analysis requires a conceptual framework, which draws on many conceptual tools. The most commonly used make for such intelligibility standards as causality, visualisability or unificatory power. Some of them were considered to be unique universal standards in traditional accounts on explanation, although they constitute only particular tools from a variety of toolkits that can be deployed in particular situations. Visualization is used as an example of such a tool that is applied even in the most abstract fields of science – as the MIT bag model from physics of hadrons shows. The familiarity developed by scientists towards the behavior of some solutions of an equation constitutes another example of such a tool. The authors point to a broader sense of the attribute qualitative, which is not reducible to the sensory or perceptual register. This sense could be documented by the meaning-shift shown by some traditional terms and they exemplify the idea through the notion of mechanism. The way it is used in modern physics (as for example in case of Higgs mechanism in quantum field theory) does not cover anymore the old meaning (found in classical physics) but it provides still a qualitative framework for scientific investigation.

Further on I will evaluate the account by pointing to its strengths and weaknesses, as well as the open questions and put them into the perspective of the goal I'm pursuing. To begin with the positive points – the fact that the account addresses the topic of understanding directly rather than suggesting a possible solution via some approach to explanation or in the shadow of such an approach, is probably the most significant one. The account does not make any use of an explanation schema; it leaves open the possibility to implement any type of explanatory scenario in promoting the goal of gaining understanding.

Another positive aspect lies in the fact that the account manages to place pragmatics in a central position. I think this is a valuable contribution to the effort of tackling the difficult issue of scientific understanding, and articulates an account by letting contextual influences play a central role. The good news from the account is that it is possible to pin down some sort of objective criteria for describing the understanding process.

The range of modalities for achieving understanding that the account intends to integrate is very broad. The difference from other approaches that advanced such an integrative view lies in the variety of means indicated to reach the goal. While Lambert-Schurz' proposal draws on different forms of explanation, this account intends to cover means of very different nature that tend to evade a general common characterization. For Schurz different explanatory formats are exhibited in the variation of the relation "Prem => Con"; meanwhile it is difficult, if not impossible, to find some common characterization among the means that promote understanding in de Regt–Dieks' account.

Turning to the weaknesses of the account, a first dissatisfaction usually voiced against any such type of account that emphasizes the pragmatics of the scientific practice, is related to the lack of explication of some of the concepts involved. As in van Fraassen's case, where the relevance relation is the notion left unexplicated (as disclosed in Salmon's and Kitcher's critique), in this account the intelligibility concept is on a similar position. Some critics, as for example Trout, expressed serious doubts regarding the usefulness of appealing to the concept of intelligibility in playing such an important role in the explication of understanding. As Trout puts it: "it [the intelligibility] doesn't sound like a propriety that much illuminates a discussion of understanding."

On the other hand, it is plausible that intelligibility presupposes already some involvement of understanding. To recognize consequences in a theory presupposes in a sense that we already have an understanding of a certain kind of it. This aspect makes the CIT criterion vulnerable to an accusation of circularity.

A second critique can draw on the fact that the criteria are too weak. To see potential counterexamples to CIT does not seem to be difficult. A scientist can take a theory to be intelligible without necessary engaging a qualitative insight. He could proceed simply by accepting the principles and laws of the theory that exhibit some plausibility and by pursuing further the consequences also making recourse to calculations. S could accept T as intelligible even if he is not able to recognize from the beginning qualitative consequences. Or, in other words, the absence of the initial qualitative insight will not necessarily qualify the theory as unintelligible. On the other side, it will be also too strong to ask for the qualitative analysis being a sufficient condition. A scientist S can see the qualitative consequences of some theoretical assumptions though it does not hold it to be intelligible under a standard that he holds on. Still less they will claim that it offers the needed understanding for the phenomena under investigation. It can develop it only with the intention to see how the assumption works, to test their limitation or for other heuristic purposes.

Another critique could point to the lack of a minimum articulation of the connection between understanding and the explanatory forms. There is practically nothing in the account that might say something about scientific explanation and how it possibly relates to the explication given on understanding. But I do not see this as a major drawback since the account seeks explicitly a direct approach on understanding and implicitly avoids this way falling on an explanation–biased route.

The following more general critique can also be considered: that the position cannot be backed by a sufficient strong argumentation without recourse to further resources that will force it to collapse on a position it tries to avoid. On one side, one could point to the fact that building understanding by extracting qualitative consequences from a theory involves drawing on many elements that would count as outside of the theory. That recalls the resources of the unificationist approach, which provides in this sense a more complete view – or something of this sort will be the strategy of an unificationist to invalidate de Regt and Dieks' proposal. On the other side some extreme pragmatist will point to the particularity of the situation to which the theory applies, and stress that the building of understanding draws on many other sources contextual determined and that the intelligibility of the theory as defined in the account does not play any overall major role. To respond to such accusations will involve the task of arguing why it is relevant to place understanding at this level and deny going in any of the directions suggested.

One could even accuse the approach of subjectivism (possible a mild one), making understanding too dependable on the scientists as individuals who assert contextually the intelligibility of a theory. In what sense is this intelligibility shared among inquirers, in other words, how do we secure a minimum objectivity? This critique can be tracked down further to the request for more clarification of the intelligibility concept.

There is a sense in which the understanding as explicated here can be taken as insufficient or lacking the needed consistency in order to be indeed the genuine understanding of P through T. The notion of understanding considered by the authors counts in this sense as partial, due to the restrictions to the pure qualitative character of the means for achieving it. The quantitative means should also contribute to the understanding of the phenomenon. The ability to use a theory comprises the skills to deploy arguments by drawing on its resources irrespective of their qualitative or quantitative nature. Only an initial qualitative insight gives often a rather superficial mode of understanding the phenomenon; the ulterior pursue of the consequences, by appealing also to quantitative means, could reveal unexpected views and reshape our understanding.

If we further take seriously the idea of models as mediating the application of a theory to the phenomena, some unwanted consequences result for the criteria. We establish the intelligibility of the theory according to CIT by recognizing qualitative consequences and this happens in the frame of a model or type of model. For a theory that intends to account for a specific kind of system, this fact is unproblematic; not so for a more general theory that comprises many different systems. It is the case of 'mature' theories as in physics, as for example the fundamental theories. They provide a more general scheme and their intelligibility is not defined in the frame of a specific model. The qualitative consequences vary with the type of system to which they are applied. For example the understanding of various other phenomena that are not gases through statistical mechanics will not appeal to the intelligibility given through the ideal gas model – as for example the magnetic systems that are accounted for through the Ising model. For a proper understanding of a new phenomenon, it will be crucial to choose the right type of model that provides the intelligibility of the theoretical frame according to the CIT criterion. What all these qualitative views have in common will usually be captured in an abstract schema that will promise us a general "qualitative feeling". The authors would like to subsume such a general type of feeling under the means to pursue the building of scientific understanding. But this does not meet my objection. My point is that you need the frame of a model to establish through CIT the intelligibility and the criterion is therefore not so unproblematic transferable when the theory applies to other kinds of systems beyond the standard intended.

Turning to some possible morals relevant for the inquiry on understanding through models we can point to the following. The account takes theories explicitly as the only entities of reference for scientific understanding. But it does not assume any specific conception for a theory or suggests constraints on how the specific features of theories contribute to the building of understanding. In fact, the extrapolation to some types of models appears quite natural. In a good sense, the CUP criterion is even better suited for models than for theories since "to have a theory for a phenomena" is better rendered as having a model (generated in the frame of that theory or not) for that phenomenon. Nevertheless, qualifying a model as intelligible would constitute a quite ambiguous task. If some condition of intelligibility in the sense of CIT could be imposed, the criteria will most probably be successfully applied in some specific sort of modeling process, i.e., one in which a qualitative insight is at work. This line can make for a working agenda to be followed in seeking to articulate a type of understanding through models.

Two further observations regarding the application of the criterion to models instead of theories. One will tend to retain in a sense the reference to the theory in the form of seeing the criteria suitable only for highly–theoretical models (the examples of models that the authors bring fall into this register). But it does not seem to be the case to state such a claim. On one side we can identify at all levels of modeling – from the phenomenological up to the highest theoretical one – such an episode of developing qualitative insight into the consequences of the model's basic assumptions. Second, there are theoretical models that are built by going the other way, i.e., by dropping the qualitative insight. We find an illustration of this point in a classical example of a theoretical model used in the explanation debate: the standard example Hempel uses, i.e., the model in social science. This kind of model is built to express in a formal way (and usually to make available some quantitative estimations) the theoretical ideas that are of qualitative nature. There is an understanding already given in the frame of a qualitative theory and the models fix it for a specific purpose as for example providing a computational algorithm for testing purposes.

To briefly conclude we can say, in general, that the account proves to be a bold proposal offering a totally different way of approaching the understanding problem. Unfortunately, it exhibits numerous problematic open points not only in its details, but also in the main lines. Engaging in an effort to fix them in the initially intended frame of application appears to be a difficult if not impossible enterprise. The account seems to have rather more chances to become successful if it limits its target on specific sorts of understanding and scientific constructs. From a general perspective, the view that the account adopts, by placing understanding in terms of the process of scientific inquiry, is also the one appropriate for addressing the understanding and explanation issues under the modeling view.

From the point of view of the basic ways of construing understanding, the account exploits the idea of understanding as ability to manipulate scientific entities. In his critical article against Trout's paper (which I will discuss in the next section), de Regt invokes Wittgenstein's claim that the understanding of a formula involves 'being able to use it' besides simply knowing it. We can interpret the criterion of successful manipulation to be in the case of de Regt and Dieks' account the qualitative insight that we gain. There is an idea among the modelists that emphasizes conceiving of understanding through a theory as the ability to engage in manipulation of scientific constructs. The other constructs that we can see as being subjected to manipulation are the models. The idea is still underdeveloped but it appears to be more promising to be pursued. De Regt and Dieks' account can make for a particularization of this idea, in the way that we do manipulate the models in order to gain a qualitative insight into the phenomena that enhances our understanding. I will address later in the text some aspects related to this conception of understanding through model manipulation.

Trout's critique of the concept of scientific understanding or the naturalists returned

In a recent paper[54] Trout voiced a bold critique in the direction of rejecting a specific sense of understanding as inadequate and even dangerous to be drawn upon in an explanation-account. Generally speaking, it is directed against some bold features grounded in the subjective side of understanding as revealed also in the specific phenomenology of understanding. His account is a naturalistic one, drawing on psychological documented studies of the two features.

The sense of understanding that Trout points to is that of a 'systematic but regrettably inaccurate index of intellectual achievement'. It is the product of these two biases: hindsight and overconfidence. They are well documented in studies of cognitive psychology. The philosophers who prioritize on the phenomenology of understanding are running the risk of contaminating their account with distorted forms of understanding that draw on these features. It constitutes a "poor cue to genuine understanding". The hindsight-bias is expressed by "I-know-it-all-along effect" in which people tend to overestimate how probable the event was before it occurred. It gives us a false understanding of an effect and makes us regard the search for an explanation as complete. Overconfidence makes us overestimate the correctness of our beliefs. As an effect, it could also prompt a stopping rule for pursuing further explanatory inquiry.

The two features mentioned above constitute without doubt pathological manifestations of explanation and understanding, but I think we have to take into account that they could become dangerous only under some conditions. In mild forms, they are a constant presence of the two cognitive processes. The problem is not if they are present, but if the kind of inquiry activity under which frame they built up inhibits the possibility of considering these biases as genuine forms of understanding. Scientific activity provides the means to do this: the way it pursues the building of objective knowledge censures the possibility of taking the above features as genuine forms of explanation and understanding.

The importance of the two features mentioned above for the process of scientific understanding is not uncontroversial as reactions showed. The suspicion rises that Trout directs his attack to a straw man position, as de Regt puts it; no author will draw on such extreme subjective features in building an account. But Trout's direct claim is that this sense of understanding is 'an important component of the orthodox theories of explanation'. Nevertheless, the claim remains much unsupported despite the author's evaluations of these theories.

Trout discusses critically the traditional accounts on explanatory understanding and finds them inadequate. Some of the critiques recall points that were made already in the debate on these approaches though they get new valences in the context of the cognitive psychology. Other critiques draw on new insights based on evaluations from cognitive psychology. But his critiques do not back the central claim of his accusation in showing how the accounts "tacitly or explicitly" appeal to the deviant features and how these biases corrupt the approach to genuine understanding.

---

[54] See "Scientific Explanation and the Sense of Understanding", in *Philosophy of Science* 69 (2), 2002.

Trout accuses the orthodox theories of explanation to be based on a wrong view of explanation as a meta-cognitive process, instead of taking it as an inarticulate activity or skill. Understanding is conceived under a 'reconstructive' view, in which the scientist pieces together information already in his possession in a detective-way pursuing a coherent story. So, Salmon's global understanding gained by seeing how the facts fit into the general scheme of the things (the world-pictures) is too unspecified and confuse; even more it is inadequate with regard to scientists' explanatory practices that prove to be local, i.e., in drawing on specific details of the situation and deferring any global explanation to outside experts. Meanwhile, according to Trout, 'when philosophers piece together such a global explanation they look at the subjective sense of fit as a cue for an acceptable explanation'. But, given the above critique, I think it remains totally unclear how the biases enter this story.

In Trout's view Friedman's connection between understanding and the reduction to a small number of fundamental independent phenomena concerns rather the cognitive efficiency and tractability of the knowledge. And further, with regard to Kitcher's approach Trout sees the internalization and identification of the argument patterns as an effect of learning and of scientific pedagogy; an 'openly causal account' could reconstitute it in terms of causal features of the internal representation of the explainer and of the relevant parts of the world. In order to support the claim, Trout points to the psychological studies that revealed the implicit learning, i.e., without awareness, which takes place in many cognitive processes.

In emphasizing understanding as a consequence of a learning activity and by pointing to the importance of the implicit learning by appeal to the tacit models, Trout indicates a totally new direction of development for the subject of explanation through models, i.e., one from a naturalist's perspective. There are some other items advanced in his suggestions the main one being that understanding is an epistemic virtue only if it is positively correlated with accurate causal descriptions. For Trout, an accurate theoretical model that gives us access to the main causal influences allows us to learn without exposure to negative instances, and provides us understanding. But he is not pursuing a causalist line in suggesting what a genuine understanding is. In his paper responding to de Regt critique, he advances some explicit[55] conditions that describe a reliable indicator of genuine understanding. To be a diagnostic indicator of genuine understanding, the sense of understanding has to be non-accidentally related to the following states of affairs: 1) the belief putatively understood is at least true; 2) the agent has sufficiently collateral theoretical knowledge relevant to the belief and 3) belief is produced by a reliable process (perceptual or cognitive). In other place Trout mentions also the requirement for understanding to be an epistemic virtue of a positive correlation with causal description of the fact understood.

Beyond these sketchy details we will have to ask in the end what morals could be drawn for an account on scientific understanding through models from Trout's arguments? A first remark is that Trout's positive attitude towards an approach to explanation through models is speaking for a stronger

[55] In his response to de Regt's critique in "Paying the price for a theory of explanation: de Regt discussion of Trout".

agreement of his exposed ideas with this kind of approach. Unfortunately, there is not much to follow in this sense. Considering his claims about the two revealed biases I think that they cannot be restated in the intended form in the context of model-based understanding in science. Although it is generally accepted that models are constructs that appeal to the subjectivity of the inquirer that imply the existence of these biases, such biases could be present in mild forms, even playing a positive role in the economy of scientific inquiry. Trout uses a specific meaning of models, referring to the tacit models that a subject develops and accompanies implicit learning. In this sense he will see models as secured cognitive products stripped away of the influences of such biases; they could have played a role in the initial stages of the modeling process, but they fade away as the model gets stronger confirmation. The reference to the correlation of understanding with the causal mechanism is one modality to guard against the biases, but it limits the intended application of the account to a particular type of model, the causal one.

In order to use Trout suggestion for approaching explanation through scientific models one has to consider the place of the tacit models which human subjects develop during the implicit learning process. We are confronted here with a gap that has to be overcome by matching the notion of models (and the results for their studies) as used by cognitive scientists with the results and insights into scientific models as understood by historians and philosophers of science.

The conditions for a reliable indicator of genuine understanding have to play a central role in removing the above gap. In principle there is no contradiction in putting together all elements, i.e., the above-mentioned conditions, the tacit model from the implicit learning and the correlation of the beliefs with causal mechanisms; but to secure a viable account on them does not appear to be straightforward. Trout’s observations are nevertheless valuable and they should for sure inform the search for a pertinent way of attacking the problem of understanding through models.

Finally, regarding the heated exchange of replies between Trout and de Regt, the main question one might ask is what is at stake in fact in this debate. Trout mentions shortly the possibility to view the divergence as between a normative versus a naturalistic approach; but he dismisses such an interpretation and accuses de Regt of adopting a rather bad strategy in argumentation by drawing selectively on some historical episodes in order to advance his thesis. For de Regt the disagreement is between the rejection through an irrelevant critique of the possibility of any account on scientific understanding versus an explicit proposal of a pragmatic approach. But in principle, we can view the disagreement between the authors as one in the ways of how to pursue the development of the subject. It is premature – due either to the underdeveloped view or to the problematic points revealed – to draw radical conclusions about which variant makes for the viable strategy to engage in. There are nevertheless points to be exploited for the accounts and lessons to be considered from the debate the parts engage in.

On scientific understanding – for a more generous perspective

Before pursuing further with the modelist approaches I will make a tour in this section to review and evaluate some more general distinctions that were made in connection to the understanding topic. These distinctions could be taken as constituting major guiding lines that one can adopt also when approaching the subject of understanding from a more generous perspective. They were usually incorporated in the basic assumptions of the accounts, a fact that underlines their importance. A brief overview of them should be very useful for setting the landmarks of the inquiry of the understanding through models.[56]

To begin with, I'll consider first the modalities in which the distinction between understanding and explanation was captured. The quest could be raised regarding the existence of a gap between the two and how it can be properly captured. A first way to reflect this will be to check for the case in which we could have one without the other, i.e., explanation without understanding and vice-versa. We have a more robust intuition on the case in which understanding could be present in the absence of explanation. It could be blamed on the lack of the means for building a proper explanation; the lack in capacity of spelling the understanding out into an "open-articulated" explanation due eventually to the particular way understanding is gained. On the other side, an explanation without understanding will exemplify a kind of un-comprehended explanation, i.e., an explanation uttered in ignorance. It will make also for an unwarranted sort of explanation. The experience of learning by imitating but not really grasping the reason behind comes to mind as a possible example. A modality of rendering this situation is by taking understanding acquired from specific sources that could harbor only one of the acts, but are improper for the other. Such case could be invoked in which understanding is gained via pictures or other non-propositional sources. But one could also conceive the case of understanding gained through propositional means that do not constitute an explanation under some adopted standards.

A bold proposal can be advanced by cutting a radical difference along the propositional versus non-propositional distinction. Explanation does involve propositions, meanwhile that seems not to be really necessary the case for understanding. It appears to be a robust intuition but it lacks an articulated formulation, and its absence from previous accounts discloses the difficulty to articulate the view. Nevertheless, there are some recent authors who adopt such a definition explicitly; for example, Linda Zagzebski states that: "understanding is the state of comprehension of non-propositional structures of reality."[57] Other authors do not seem to engage in such a radical viewpoint but search for a more qualified one. For Kvavig propositional understanding is a specific brand of understanding that we refer to when we attribute understanding in form of a propositional operator. But he lets open the possibility that not a set of singular propositions but rather an "information chunk" is the object of understanding. No further suggestions are given of how we should conceive this 'chunk'. Of course, both types of

[56] In his 2005 conference on *Philosophical Perspectives on Scientific Understanding* P. Lipton delivered an overview of such main intuitions that are at work in the field, and I'll partially draw in my exposition on his talk.

[57] in *Recovering Understanding* p. 242.

proposals open a rich agenda of issues to be clarified. For some tastes, especially the ones probably from the philosophy of science (the previous mentioned distinctions were proposed by epistemologists), it's possible that trying to cut out a difference on his line, loads the problem with a too heavy philosophical burden. This would explain why not too much about such kind of distinctions is to be found in topics of research in philosophy of science.

A different sort of proposal might try to modulate the explanation–understanding distinctions along separation lines already invoked in the explanation debate. Such one is the actual-potential distinction. It was applied to explanation in order to capture a tentative explanation from one that was already confirmed. Understanding as discussed till now in the explanation approaches was taken rather (though not explicitly stated in the literature) as referring to the actual type of explanation. In Kitcher's or Salmon's accounts the understanding referred to is one given through the successful explanations. As P. Lipton noticed, we could consider cases of (actual) understanding from a merely potential explanation. One could ask why this was not done already. One of the reasons that can lie behind this fact is that an understanding, given by a potential explanation, could be considered as being rather preliminary and unreliable. Moreover, extending the qualification actual-potential to understanding raises new issues even if we could parallel the interpretation from explanation in case of understanding – meaning an actual understanding will be considered a genuinely confirmed one or a fully objectified one (to put it this way). But what will a potential understanding from an actual explanation or vice-versa mean? This opens some further issues, and how to handle them is not obvious. Anyway the distinction is not one that will cut to the bottom the difference understanding – explanation, but it may add finer local qualifications.

A characteristic that signals an important difference between explanation and understanding is to be found by viewing explanation as being rather a public propriety, while denying it (at least partially) for understanding. A qualification of some sort is involved since we would like to attribute objectivity to scientific understanding. But scientific understanding is public not in the same sense as scientific explanation. To distinguish properly between the sorts of publicity conferred to explanation and understanding we do need in the first place a developed account of what they are.

Another way of distinguishing between the two implicitly assumed in explanation accounts construes the two along the lines of the distinction between process and product. The usual forms for this distinction involve taking the goal (or the main benefit or the main aim) of an explanation to be the generation of understanding. In this sense, understanding relates to explanation as the product to the process of production. The drawback of this strategy from the perspective of a direct approach on understanding lies in the fact that it pushes us back into an explanation-centered way of dealing with the problem. The explanation itself is also subjected to the process-product duality. As a process it refers to the very act of unfolding an explanatory story, while as a product it refers to this story directly.

But, what will be to consider a process-product distinction for understanding? Understanding is usually conceived as a state, which will contradict taking it as a process. Nevertheless, by insisting on the idea of understanding as a process we might view it analogous to explanatory acts: the act of understanding that will take place during the inquiry process. An understanding as a final product would find a good place at a higher-level of scientific activity: as linked to the final state of cognitive enterprise as expressed through a stable scientific construct. Models can recuperate both senses: the process of understanding as reflected by the specific modeling processes but also as a product of scientific inquiry.

From yet another perspective to have understanding it means to be able to produce such explanatory acts, to reenact the right explanatory story inducing understanding in the receiver's mind. Explanation is placed under this view in the position of a mediator through which understanding is transmitted, and enriched. Though Lambert and Schurz take the basic construal of the relation between explanation and understanding to be embodied in the act of yielding of understanding through an explanation, from their account we might also read out a view on explanation as a mediator between two different states of understanding.

As to a direct characterization of understanding, an angle to tackle the problem makes recourse to the phenomenological aspect of understanding. The quest that would first drive our concerns under this aspect will refer to the specificity of this phenomenology. The Aha-Erlebnis is the standard symptom that is usually referred to in this case. Suspicions rise in an analogous way to Hume's reaction to the phenomenology of causal links, i.e., given through the feeling of expectation. The possibility to have the feeling without really having understanding or vice-versa raises doubts as to how much one has to appeal to phenomenological aspects. Trout's critique explicitly points to this aspect, as discussed, but no all authors will agree if some characteristics that understanding exhibits should be consider as phenomenological manifestations only. The boldest example in this sense is coherence and its associated 'feeling'.

Last but not least we have to mention a construal that appears to be most promising, i.e., the one that identifies understanding (totally or at least partially but in an essential way) with a kind of ability. Understanding as ability could be seen as a kind of criteria of externalization, of objectivation – avoiding making understanding a too internal, subjective affair. It captures also the pragmatic and dynamic aspects of understanding. The necessity appears to be bolder than the sufficiency of the ability as a condition for understanding. An understanding without the ability to deploy it – if this means giving explanation or involves also other manipulative abilities – makes for an incomplete sort of understanding. Zagzebski draws on specific interpretations[58] of Plato's texts and advances the idea of taking the concept of episteme as referring rather to understanding than to knowledge and further conceiving it as being acquired through the process of learning an art or a skill. Understanding is acquired through mastering a *techne* even for academic areas and so it follows that ability is built-in.

*** * ***

58 Especially those of Moravscik and Gail Fine.

As the previous references showed, epistemology is an area of philosophical research in which the concept of understanding was invoked in recent debates. The orientation known as virtue epistemology is the one that harbors such initiatives. The main reason behind the appeal is to redefine old problems in epistemology by advancing new ways of approaching them. Kvanvig argues, for example, for a solution to the epistemological problem in Meno, by taking understanding instead of knowledge as reference. For Zagzebski only the skeptical periods were dominated by the justification-certainty topics, while the askeptical ones took understanding as the subject of their concern. W. Riggs for example argues for the analogy of understanding to *eudaimonia* as an aim of exercising the intellectual virtues. All these authors emphasize the value of understanding and call for a reconsideration of epistemological issues in the new light.

An ongoing debate dwells around the issue of the differences between understanding and knowledge. Authors as Kvanvig and Zagzebski argue against understanding being a sort of knowledge, while Grimm tends to see understanding as sharing more aspects with knowledge, though not being identifiable with it. The separation between the concepts is drawn through such facts as the one that knowledge requires, truth and exactness, but understanding does not (Zagzebski); or that understanding is transparent, i.e., consciously transparent, as Zagzebski states: "understanding is a state in which I am directly aware of the object of my understanding", but knowledge is not; or that knowledge is incompatible with a kind of luck while understanding is (Kvavig). To play down the sharpness of the distinctions makes for the opposite position.

A common locus is to take understanding as being constituted of knowledge enriched with some extra-ingredient; some authors identify this ingredient to be given by the capacity of grasping the coherence-making relations among beliefs or as they are also referred to (for authors like Grimm or Lipton) as the capacity to grasp the explanations. Grasping is the term used to cast some clarification light on understanding but it is not sufficient since it rather paraphrases it. Grimm is the only author who engages in an effort to articulate what grasping should be in an analysis of first person inquiry experience. The grasping is rendered in terms of a "fallible exercise of his [one's] capacity for seeing dependencies" in which the cognitive subject is not only "passively taking he world as it unfolds before him" but is reaching out towards understanding the dependency in a new and distinctive way'.

Leaving aside the above details, which will recur in some further different places, the pertinent question we can ask, from the point of view of our agenda, concerns the kind of contribution such inquiries can bring to our project. First, it is to be noticed that very little was done in the frame of this debate to provide an articulation of understanding as a main objective of an account. There is no account comparable with Kitcher's, Lambert-Schurz or de Regt to focus direct on an explication of understanding. The concept is rather tackled and proposed for naturalization in the epistemological field under promises of better perspectives on old epistemological problems. The main effort is directed therefore to accommodate the notion relative to the other epistemological concepts and argue for the advantages it can bring by taking it seriously. This undertaking is accompanied by valuable characterizations and distinctions on the understanding notion to be taken into consideration in the context of our inquiry. It might also constitute that relevant advancement in articulating a satisfactory theory of scientific understanding through models have to wait for more consistent results in the epistemological field; but given the actual situation there are many good insights coming from epistemology that might provide us with valuable suggestions in our investigation.

Modelist approaches

Now to return to the modelist approaches (or model-centered ones) in this section I'll browse the landscape and pick out the moments in the literature touching in a certain way the subject of scientific understanding. I've already mentioned some authors who had already something to say on explanation in the models' context; unfortunately little is said on understanding. The area of research seems to be open but rather unsettled and quite incipient and meager in results. As the Amsterdam 2005 conference showed, there is a strong interest to deal with these issues but the huge difficulties of the subject make for the scarcity of viable results. I'll evaluate the issues under the main directions of approach from the modeling literature and point to the most appropriate ways of conceiving understanding in different frames.

The pragmatic approach on models (as explicitly stated by M. Morrison) enjoyed much of the attention during the last period as constituting a main direction of development of the model-oriented approach. The purposes and uses of the models are the central aspects through which they are approached. In as far as it concerns the subject of scientific understanding, there is no systematic analysis undertaken in this frame. Nevertheless, the explanation topic received some attention, though it was not intended to make for an articulated account on the subject – I already discussed in the previous chapter Morrison's claims about the explanatory qualities of models – explanation is seen as a product of the representational function of the models. Given the close connection between the two topics, we could expect that scientific understanding has a similar status that places it in the register of the representation function.

M. Morrison does mention in her programmatic papers the fact that we gain understanding through models. In her example referring to Prandl's model, she asserts that the water tank provided some sort of conceptual understanding for the building of the mathematical model that could not be gained directly from the theory or on experimental ground. She does not give us a characterization of what such an understanding is, and makes use of the term in a more general and informal way. In her further claims about explanatory qualities of models, no further clues are given to illuminate the link between explanation and understanding. Besides this open issue, Morrison's account discloses the need for accounting on how different models relate using the understanding gained in one model to produce understanding in another one. We can read out here some indirect suggestions as to the existence of different levels on which understanding of the phenomena builds.

There is some more relevant information for taking understanding as a manipulatory capacity in what M. Morgan reveals regarding the model-based reasoning. In her paper *Model Experiments and Models in Experiment* she identifies four distinct moments in the process of reasoning through models: 1) the building of the model; 2) the raising of questions together with; 3) the manipulation of the models to extract answers to questions, and 4) the step of connecting these answers to the real system. Manipulating the model to produce answers to the inquiry questions is only a moment among others that characterize the modeling process. A specific kind of understanding, as the ability to manipulate the model, will build up only when we engage the resources of the model to answer the inquiry questions. This could be taken as a too particular kind of understanding in order to make for a sufficiently adequate candidate for a more general model-based understanding.

In building a model, one engages in manipulating concepts and scientific constructs that possibly were developed previously in the frame of theories for other models. This discloses other types of manipulations and an understanding that builds at the level of theories or even entire fields of science, since there are no constraints set on the appealed resources. It can be viewed as a given pre-understanding relative to the phenomena targeted in the modeling process. In fact, every step from the Morgan's list discloses some understanding that can be conceived as manipulation since it involves the use of scientific concepts. The question that arises is whether they are moments contributing to a more complete and unique understanding or they make for separate types of understanding. In this later case, the main problem to be asked is which one will be the most suited to be identified with the understanding delivered through a model.

Some bold questions independent of the context of inquiry that rise in connection with the construal on understanding as ability to manipulate require urgent answers in order to secure a clear frame of investigation. Naturally, we have to ask what the subject of manipulation is. We claim understanding through some entities by manipulating them or the constitutive parts of them. We gain understanding through theories by using them in the scientific enterprise or by using parts certified by belonging to some theories or we could require the use of some essential constituents in them. A correlated problem with the previous question asks how does the nature of the entities that are subjected to manipulation influence the sort of understanding that we gain. These are issues related primarily to the objects that are to be manipulated. Other important issues are generated by the fact that we have to qualify properly the manipulative act. We cannot accept any kind of manipulation, except the one that makes the construct workable, i.e., it is a successful manipulation. How are we to characterize in an adequate way such a qualification? A subsequently related question is if we could justify the adequacy of the criteria. Such questions do need to be addressed to give some real consistency to the idea of understanding as manipulatory capacity.

Before ending this section I'll address a reaction that may arise in relation to my discussion above. To put it directly some would ask: why should we charge the pragmatic view on modeling with the understanding as ability conception. Where does the naturalness of this charge come from? In fact, the authors from this camp opt for the representational approach for the explanation problem. But the naturalness of reading an ability construal for understanding is directly justified by the central intention of the pragmatic approach to emphasize the functionalist view on models. We can also interpret that the reason can lie exactly in the difficulties to implement the construal of understanding as ability, which makes the adherents deliver the problem of explanation (and understanding eventually) to the representational approach.

A short mention should be added also about another sort of approach to models that favors the ability construal of understanding. The authors call it an objectual approach and its attitude towards models is characterized by viewing them as epistemic artifacts.59 Their position is anti-representationist, and they plead for moving the focus from representation to the mediation and production aspects of models. The position agrees partially with the functionalist view, but its specific emphasis is on the materiality (not in a simple sense of a material object but having the possibility to materialize in different ways) and objectuality of the models. Regarding scientific understanding60 their position will fall mostly under the ability construal; for them the understanding arises in "specific subject-object relationships in which the artifact is embedded". Understanding is coupled with different uses and the understanding providing features of models is tied to the specific skills and experiences of the scientists and to the situations in which models are built and used.

The authors do not advance any general conception of understanding but it is clear that they would like a more detailed approach to the social and practice contexts of the use of models. They exemplify their claims with episodes from particle physics (Merz) or linguistic research field (Knuutila). A valuable point, which they emphasize, takes understanding not as a result or end point but as an act that accompanies the modeling process. Their position is contrary to the position that promotes understanding as a product of explanation and tends to view the two acts in more generous perspectives on scientific inquiry. The drawback of their position on understanding is that the consequences that could be drawn through their approach can also be obtained through a more general pragmatic approach.

Representational approach

Now turning to the approaches on modeling topic that focus on the representation subject, we notice that they find an articulated expression in more accounts – as the one of Hughes or Hartmann and Frigg that will be discussed further. The accounts try to provide a unifying approach to modeling through the representation notion, but the range of application is limited to the theoretical types of models. There is no incompatibility between this approach and the functionalist view of the pragmatic conception; from the last perspective it can be taken as a specific inquiry to clarify the particular

59 They draw on Rheinberger's work, where the concept is introduced.

60 T. Knuutila and M. Merz attacked the subject in the 2005 Amsterdam Conference.

function of representation of the models. Probably many of the authors that work under the representational agenda will also agree in broad lines with the pragmatic theses.

Hughes is the author who provided us with an explicit articulation of modeling process through the representational view in his papers *Theoretical Explanation* and *Models and representation*. In the first paper he addresses explicitly the subject of understanding but no reference to it is made in the last one. That is partially justified since the goal of the first paper was to account for a type of explanation: the theoretical one in physics.

Hughes restates as a starting point of his discussion on understanding the older idea of seeing theoretical explanation as a metaphorical description of the system under investigation in accounting for the cognitive function of models. But Hughes proceeds by stressing the differences between the metaphors and models that inhibit a further development of the positive analogy; he draws some conclusions based on the divergences revealed. In this sense metaphors require that the second system – in terms of which the re-description is made – should exhibit familiarity to us and provide this way 'a rich store of associations' for understanding the first system. Not so in the case of a model which is not a better-known 'pre-existing entity' but a 'free creation' that might appear unfamiliar. The quest is then how does the model acquire its cognitive function and, in a bolder formulation, what takes the place of the pool of associations provided by the second system in the case of an abstract model? The two suggestions he makes in order to clarify this issue are: first, the fact that the resources of the model are discovered gradually through investigation into it; second, that an abstract model has a greater volume of associations than any concrete system (associations that are gradually revealed). It is directly inferred that only putting the model to use we build the referred understanding.

But none of these two suggestions tells us directly what understanding is or how is it gained through a model (besides the fact that it is the understanding provided by familiarity with the second system). What one may infer and it is not explicitly stated is that in comparison to the metaphor where one system is understood by translating it into an already exiting configuration of things, in the case of the model we do engage in an act of building or discovering of such a configuration. These actions are reflected in Hughes' account on theoretical explanation through the construction and the act of demonstration that takes place in the model. His account states that: "a feature X of the world is explained by displaying a model of the part of the world and demonstrating that there is a feature Y of the model that corresponds to X and is not explicit in the definition of M". The internal dynamic of the model discloses the process of explanation.

Hughes does not advance any explicit proposal on scientific understanding; he makes rather scattered remarks and suggestions drawing on well known positions on understanding. His appeal to the metaphorical re-description conception does not intend to induce any metaphysical agenda (as he explicitly specifies) but is only meant to draw "a parallel between linguistic and scientific understanding". This draws actually on the idea of understanding as ability or competence to rightly perform the linguistic or explanatory tasks; thus in the case of scientific understanding to "follow explanations in terms of a theory […], paraphrase these explanations and to find analogues for the behavior of the models." But this refers directly to understanding given through the theories and not through models.

From the text transpires also Hughes' inclination for an unificationist solution on understanding approach but besides a brief mention, he does not pursue systematically such an articulation. This restates his choice to take scientific understanding primarily as given through theories or larger units (such as fields or entire corpuses of knowledge).

Hughes' suggestions on how understanding could be seen are even more various than the above-mentioned claims. In other places, he refers to understanding as allowing us to become aware of the representation we use. This draws us into a more subjectivist register but recalls also another general view that conceives understanding as the capacity to grasp explanations. The last claim didn't find an adequate development, but is still used in an informal way and remains problematic (the claim makes also for a general guiding intuition under which an account could be articulated). A short end notice that could be worth mentioning is the idea of a model as an 'epistemic source' that is exploited. The suggestions coming from the epistemological insights into understanding are to be implemented under this very generous idea.

*** * ***

In a paper on the model-based understanding in condensed matter physics, Yi addresses the issue of understanding through theoretical models in physics. No explicit model of such an understanding is proposed, instead some distinctions and observations are advanced in how we should approach the topic. He begins by eliminating two candidate positions and points to the plausibility for a third one. The rejected positions are van Fraassen's empirical adequacy and Hacking's 'instrumental utility'. Empirical adequacy is an extra virtue of a model reflected in the predictive success that does not make for the understanding provided by the model. No discovery of a novel effect will induce understanding through the model as in Hacking's criterion. Yi prefers Hasok Chang's proposal in which the pursuit of ontological plausibility in our system of knowledge is the condition of intelligibility for our scientific theory.

But Yi's contribution is to point to a distinction that we have to take seriously: we should be able to understand the model before claiming understanding through the model. The understanding of the model consists of three main steps: first – determining the consequences in the model about the target system; second – figuring out a physical mechanism, and third – developing our physical intuitions about the model. According to these steps, we explore and extract the consequences about the system by deploying mathematical techniques, then we figure out a physical mechanism by making reference to an interpretative model, i.e., a concrete implementation of the abstract structure, and as the last step a set of "canonical intuitions" is made possible when we get at least one successful application. Only when these steps are completed we can engage in understanding the target system through the act of matching one of the interpretative models with the central features of the phenomenon. All in all the above descriptions define his two stage model involving, in a first stage, the building of an understanding of the model and in a second, the matching of a target phenomenon with a well-motivated interpretative model.

Yi stresses the difference between the understanding of the model and the one of the target system but, I think, this is of value only for such a particular case as the one referred in the account, i.e., one that involves two distinct levels of modeling – a pure mathematical one and an interpretative one. I'll discuss later how this difference can be recast in a more general way in Hartmann-Frigg's account on explanation. But even restricting us to the intended model, the main objective of clarifying the nature of the model-based understanding remains partly untouched. In fact the problem shifts only to yet another one. Since the 'representational value' of such an abstract model lies 'almost entirely with the successful interpretative models', the problem is further restated by asking how do the interpretative models provide understanding.

Now, related to the model-based criteria advocated by Yi, i.e., "the pursuit of ontological plausibility", we could also accuse the fact that the understanding it points to is induced rather by theories at the model level than by being built in the models. This happens because the development of physical intuitions, which play an important role in Yi's proposal, takes place in the frame of theories rather than models; or to put it in terms of Hasok Chang's ontological plausibility, the idea that Yi makes reference to, the plausibility is given through quite general principles which reign over the entire corpus of knowledge.

In the end, taking a bit distance, we should ask what does actually Yi's distinction between the discussed understandings signal. Is it a sort of distinction that can be viewed in an analogous way to the type-token one? Or it points only to a distinction between some abstract, i.e., more theoretical models and some more concrete ones. This last one can be restated under the perspective of a more principle-guided or theory-driven model building versus a phenomenon-driven way. But this recasting seems not to be very rewarding. As I mentioned above and I'm restating here, the claim that what is obtained by this distinction is only a shift of the problem without any suggestion for a solution for the issue of scientific understanding. The move might have been productive if it would have delivered a consistent way of how a kind of understanding builds from another. But there is not such a proposal to be found in Yi's paper.

Hartmann-Frigg's account - a more general proposal on explanation through models

In their contribution at the conference[61] on scientific understanding, Frigg and Hartmann addressed directly the subject of understanding through models and in a most general way. Their proposal intended to 'make the first steps towards a model-based account' on understanding and

[61] *Philosophical Perspectives on Scientific Understanding*, Amsterdam, August 2005.

explanation without restricting them to a specific type of model. They present a sketchy[62] account mainly on explanation and advance claims by generalizing some already existing ideas in the modelists' community. Much has to be articulated to give a real weight to the proposal (and the authors are cautious and guard against taking it as a worked out version); nevertheless a discussion around the main lines can and should be done.

The frame in which their approach takes place is one of the representational view: ‘our take on explanation essentially depends on the model’s representational function’. But the authors explicitly specify that there is no need for a worked out theory of representation, nor a particular take on the notion of representation in order to articulate the account. They insist on this point although, in my opinion, they do not bring sufficient reasons to back it properly. From a certain perspective, they seem to be justified in insisting that no specific articulation of a theory of representation should affect the account. The sketch looks this way but it remains a challenge to get a worked out version given the authors’ inclination for the representation subject. There are some special claims made by the authors – some considered central – in need to be adjusted with the above affirmation of independence. One is the bold identification claim of different modes of understanding with different sorts of representation. Another claim is related to the description of the last step in the explanatory loop. Besides, further clarification of the explanatory mechanism seems to need an appeal to more detailed aspects of representation.

In the first step, they proceed by identifying the *explananda* and the *explanandum* (setting also in this way the limits of their account) in the context of model-based explanation. The *explanandum* is singled out as an ‘occurrence’ O that refers to any kind of part, aspect, feature or propriety of the target system T that is represented by the model M. The *explanans* will be the model M itself.

We encountered already in the tradition of the debate on explanation a variety of positions related to the nature the *explanans* and the *explanadum*. It is in a way debatable in what sense the authors manage to put aside the worries related to this issue. The technical term ‘occurrence’ is chosen ‘by convention’ to refer to all sorts of entities that can stay in the *explanans* position; but the request for further clarification is not entirely unjustified. The move here is, on one side, rather a delimitative one: as the authors specify the intended kind of explanation to account for is the one from empirical sciences disregarding the one in mathematical sciences. On the other side, the move discloses also the intention of deferring the issue. But as I will discuss further there are many points in the account that have to be clarified in connection to this problem.

Turning to the specific articulation of the account, which they call the “loop account” of explanation, the authors present it as being constituted of four main steps: two identification steps, and two explanation steps. The first two steps identify the occurrence in the target (OIT) and respectively in the model (OIM); the first explanation step ‘reproduces OIM in the model, i.e., it shows that OIM follows from the proprieties of the objects in the model and is not put there ‘by hand’; in the last step

[62] An unofficial draft outline of the main schema is available from the authors; it is the only written version which constituted the basis of their presentation.

we carry out this knowledge from the model to the real world by exploiting the representation relation. To provide an explanation amounts to running through an entire loop of this kind.

A first reaction concerns the way we should understand the identification processes. The identification in the target is an issue closely related to the 'occurrence' one. The second identification in the model is, on the other side, closer connected to the process mentioned in the third step: often the identification takes place as a consequence of a reproduction in the model and does not arise in isolation. Both identification processes are in need of further clarification, and to be placed in the context of other processes that take place in a model and outside a model. These processes could pertinently be addressed by drawing on the representational practice in science; but the authors' intention is to refrain from such an appeal that will require an account of representation. The first kind of identification – outside the model involves different representations, independent from the model, though they can be suggested by it. We think of experimental representations of data (as for example diagrams) that are involved at this level. In a semantic view approach we will talk of models of data or of experiments. The OIM will involve representation directly linked and generated by using the resources of the *explanans*-model. A good direction to look upon in order to search for the above mentioned clarification may involve the engagement of a well-articulated account on scientific knowledge as the structuralist version of semantic view provides. Of course, a rich agenda is required to be worked out in this case.

What the authors are doing can be viewed also as an effort to account for the construction of *explanadum*. The *explanans*, i.e., the model is assumed to be given. This raises the same objection that I formulated in a previous chapter. Model modification and model building are processes that should be considered when dealing with explanation in the models. Taking models as ready-made constructs brings back the view of explanation as application that I accused in the first chapter to be the view that misguides the inquiry into explanation problem. Nevertheless, it seems that the objection does not strike so hard since there are possibilities to accommodate in the account the model building process.

Regarding the last step of the LOOP move, which is defined through the transfer of knowledge from the model to the target system, the authors place a strong emphasis viewing it almost as a definitory moment for the entire understanding through models. In the contribution[63] to the conference they state: "it is at this point that we gain understanding." I am inclined to think that this emphasis is wrongly placed and tends to give us a reading that does reduce the model-based understanding to this moment (though such a claim is not explicitly pursued by the authors but only suggested).[64]

On one side, by reducing understanding provided by models to an understanding that builds up at the moment of knowledge transfer will match the view of understanding as the final product of an explanation process. But this fact does not justify the reduction. To restrict the meaning of understanding through models only to this moment, it will be a too narrow way of explicating it. My suggestion is that the explication has to take into consideration the other moments of model-based

---

63 A booklet of the Conference *Philosophical Perspectives on Scientific Understanding*.

64 In the written version of the contribution the claim is presented quite smoothly.

understanding that contribute substantially to the process. One objection could point to the fact that it is more understanding about transfer involved than about the target system. The understanding of the target system is not a radically new one that builds at this moment, but we engage in an adjustment of the understanding gained through models building and manipulation. Of course it can involve a specific ability to manipulate models or other scientific entities, but it is not clear that this is totally different from the other moments of understanding. When referring to this adjustment one has to take also into consideration constraints that come from representations outside the model. Besides, what is specific to this moment is a kind of declaration of the validity of understanding. This could be viewed in analogy with Giere's idea of advancing a theoretical hypothesis by which the scientist asserts that the model applies to the system. Therefore there are more considerations to be taken into account when dealing with the moment of transfer, considerations that do qualify the claim that this is the moment when we gain understanding of the system through a model.

Moreover what is to be explained according to the account refers to an aspect or part or some feature propriety of the target system. The depiction of the target system that the model offers does not provide an explanation and an understanding of the target system but of some parts or aspects of it. Some more argumentation is needed in order to establish that this sort of understanding is the one to be identified as the model-based understanding; or, in case of another option: that it is not the case to look further for such a kind of understanding, since this is all what we can get.

A particular kind of argument, which might back the above claim, would show that we could reproduce the occurrence in the model and accomplish the explanation steps according to the account but nevertheless the sense in which the model explains the phenomenon fails to be justified. I think we could exemplify the above claim drawing especially on areas of application of models that were not intended when the models were originally conceived or under the basic assumptions that guided their initial design. Such areas are to be found especially in the interdisciplinary fields of research. The account will make such a model 'explanatory' for a system on which we apply the model for heuristic reasons or other sort of considerations than the straight explanatory ones.

To exemplify the above claim I'll draw on the Potts-model - a variant of the well-known Ising model - that has its original application field in the area of physical systems and was designed to explain specific ferromagnetic phenomena. A clustering algorithm was developed starting from the model and applied to genetic data:[65] genes are grouped this way according to their activation, gaining insight into their collective behavior. We will not say that Potts-model explains or offers us a genuine understanding of the genes and their behavior, although the Hartmann-Frigg's account will not speak clearly against it. The objection from their side comes probably under the idea that OIM should not be 'put by hand' but deduced from the basic assumptions of the model. But this general requirement does not seem to be sufficient enough in order to lift the potential misinterpretation.

What's missing here? One could see a need for more constraints on the correspondence relation, which should make the process fall into the explanatory register. There are no requirements demanding

[65] The algorithm was developed by Blatt, Wiseman and Domany.

a correlation between how the OIT and the OIM is built. We could also say, following Yi's view on the model-based understanding, that an inadequate interpretative model is forced on the theoretical model. What would strike us as being quite divergent is the different theoretical background on which the description outside the model – a genetic-biological one – and the one of the models are built. One solution could be Yi's requirement that draws on the development of some physical intuitions as an important component of the understanding through theoretical models; this seems to do the needed job of blocking such wrongly targeted systems.

A modality that does not necessary rescue the initial objective of the account but rather retains its initial intention would be to view it rather as proposing a sufficiently general schema that has to be filled with more specific details in order to deliver the intended results**.** The authors' intention might have been this as it is in the case of the appeal to the representation concept. But as the above-discussed objections showed even so the schema is in need of more constraints or at least a reinterpretation that specifies that it accounts for more types of models than only the cases of explanatory models.

A way of throwing some light on the proposal is to draw a comparison with Hughes' account. Hughes' account provides us, besides the proposed schema, with certain suggestions on how to engage in further clarification of some of the concepts used in the account, as for example: the concepts of denotation or demonstration. No such options are available in Hartmann-Frigg's proposal and, in this sense Hughes' approach appears to be better backed by results from other areas of research. The concepts deployed by Hughes are in a sense already technical terms which were developed in the philosophical literature in various areas, which is not the case with Hartmann-Frigg's account. The notions of identification or explanation used by the authors in their exposition could rather obscure than illuminate their construction. Such concepts are themselves in need of explication before engaging them in the account.

While drawing further on the comparison with Hughes' account, one recognizes some of the 'identification steps' from Hartmann-Frigg's account in the denotation process, which Hughes describes in his account. Furthermore the authors mention explicitly the problem of the correspondence between the two occurrences, which should be addressed more carefully. There is a rather bold question to raise: how consistent is the perspective gained by renouncing such a proposal as Hughes' which appears to be well-anchored in a philosophical background for a totally new view that doesn't seem to get support from a rich additional conceptual background.

I would like to add here an observation regarding a fundamental attitude expressed through the following claim: "it seems that there is a different mode of understanding associated with every representational strategy."[66] It parallels Morrison's claim, which refers to explanation instead of understanding through models. In Morrison's account it is coupled with another conclusion regarding the possibility of a further clarification of the representation since, in her own words, there is 'no one way to characterize the nature of that representation'. Hartmann and Frigg do not take such a position

[66] A reference extracted from their talk at the 2005 Conference in Amsterdam. (See the abstracts).

and of course this happens because they want to defer any reference to issues involving the topic of representation.

The claim points anyway to a closer inquiry into these strategies of representation in order to gain more insight on different understanding modes. When taking the strategy seriously, the advantage is the partitioning and deferring of the hard problem to more local strategies. For some authors, Morrison’s claim might appear to be a bit too deflationary since no further suggestion is given on how to identify these strategies. Hartmann and Frigg go a step further by drawing on different kinds of representation and by providing a taxonomy of them comprising such representations as (among other) pictures, simplifications, approximations and idealizations. But this claim is not really substantiated and it actually raises problems related to the way in which to articulate the specificity of the understanding for each representational mode.

Last but not least a short comment is in place on the choice of the example the authors chose to work with. The authors take care in justifying the choice of the model on which they exemplify their account: Boltzman gas model. The justification makes sense in the context of the debate on explanation and understanding since it makes for a clear reference. The further necessary step is an exemplification on a model that is not a theoretical model. This will speak for the generality of the account.

To conclude briefly on the tentative proposal that Hartmann and Frigg advanced, it is to be noticed that the account opens clearly a Pandora box of problems but it does raise also the promise for a general and flexible frame in which the problem of explanation and understanding through models could be placed.

Despite of its drawbacks (which could be partially justified by the fact that it is only a simple sketch of an approach) its main merits lie with the fact that it really dares to address the question of understanding and explanation through models in the most general and direct way. I do think that the proposal strikes at the relevant level at which a solution should be sought.

*** * ***

Some concluding remarks are in place. In this chapter I have tried to mark and reveal the main ways in which understanding was approached in different accounts with an eye on the prospects of applying these strategies to model-based understanding. From the different approaches I reviewed, I tried to extract the pertinent points which could support a naturalization process in the modeling context. I have tried also to isolate the hindrances that would inhibit an advancement of the project.

As a consequence of our overview some main conclusions and re-orientations could be drawn. First, in comparison to the explanation problem there is no main fixed, well-coined repertoire on the subject of understanding. Of course, under close scrutiny we could detect some trends, but nothing compares with the points heavily debated in the explanation topic. One historical fact that could explain this is the absence of any account that would be similar to Hempel’s standard model on explanation. Friedman and Kitcher’s accounts cannot be seen as having the same status with regard to scientific understanding. This situation does not constitute necessary a disadvantage for our investigation – it should be taken rather as an advantage in the investigation of the scientific understanding subject.

Secondly the most elaborate accounts make a clear reference to other bearers of scientific knowledge than models especially to theories or to corpuses of knowledge. One of our first moves is to engage in tackling our topic by sorting out the relevant points in order to settle the problem in the modeling context. I've tried to suggest the ways to accomplish this move under the pertinent constraints that could be revealed given the today's landscape of the philosophy of science.

By discussing the way in which understanding was conceived in the classical accounts on explanation, my intention was to underline the fact that the construal contains some fundamental assumptions which makes it unsuitable for an application in the model case. This does not imply that valuable suggestions could not be extracted from these approaches and that some particular model types would not allow some parallel with these accounts. It is rather that their limitation is to be situated in their general lines that make for the guiding attitude to the inquiry of scientific understanding. The emphasis on the global character of the understanding, on the unificatory virtues, or exclusively on the coherence construal make these accounts to be opposite in spirit to approaching adequately understanding through models. In contradistinction to the older approaches some of the new attempts that deal with the concept of understanding proved to be more flexible in tolerating an eventual implementation in a modeling frame.

As the necessity of a pragmatic approach became clear, de Regt and Dieks' account clearly shows that such an account will fit better a modeling view. As the critique on their proposal and the further discussion revealed, the two main moves that have to be undertaken for a successful articulation of such a pragmatic approach will be on one side to take into consideration some essential constraints and on the other, to appropriate some valuable new insights, ignored till now. In this sense we have to consider as essentials the restriction given through the qualification according to the models' typology; in regard to the second move, a better articulation of a pragmatic view has to draw on a more generous perspective paying attention to the advancements in other related philosophical fields as that from virtue epistemology and their results on understanding and its relation to knowledge. I discussed also a series of fundamental distinctions comprising some basic and powerful intuitions which have to be considered and decided upon, when engaging in the inquiry effort in order to decide on the viable guiding lines of the approach.

Further on, the recent trends towards what is to be taken as the central features of understanding shifted the accents in the direction of a partial disregard of the till now heavily emphasized coherence aspect. Trout's critique radicalized the attitude by not only ignoring the characteristic as other authors usually did, but also accusing it as the main hindrance to any positive results of an inquiry into the understanding concept. A development under a naturalist conception could make this rejection more substantial but the other orientations seem not (yet) to back explicitly such a position. On the contrary, in the epistemological positions the coherence aspect enjoys a central position in the construal of understanding. This fact discloses a divergence in the tendencies in philosophical inquiry on understanding in the two fields despite their apparent similarity in working agenda.

Regarding the modelist camp, there are important points made in their accounts that had to be considered especially for specific types of models in particular scientific areas (where we could hope to get completely articulated accounts). One important quest is therefore the possibility of a transfer of different characteristics identified in such cases to other types of models or even generalization for a broader range of types. A plausible and flexible more general frame to situate claims about explanation through models is the Hartmann-Frigg proposal. A pertinent immediate working agenda would require contextualizing their general schema for specific models in different fields.

In what regards the understanding concept, it appears that construing it as an ability to manipulate theoretical constructs is the usual choice of the modelists (more or less explicitly claimed). Nevertheless even in the most explicitly articulated accounts a clearer fixation of the surrounding issues is required. In another order of ideas, one of the main conclusions that can be drawn is the existence of more moments of understanding that are to be identified in a modeling process. The issues to be addressed immediately pertain to the criteria of identification and more general to the ways of articulating properly these moments. The frontal quest addresses the possibility of declaring model-based understanding either by indicating one of these moments or showing how more moments contribute to the emergence of the overall unique understanding.

One auxiliary point to be raised in the discussion of the above-mentioned issues is to guard against overdistinctions and identification of spurious moments. Such a distinction as Yi's between understanding of the model and the one of the target system could be of this kind. A better rendering of the distinction is viewing it through the moment of knowledge-transfer as made explicit for example in Hartmann-Frigg LOOP account. However we engage in another sort of epistemic act by claiming understanding of the target system gained through our model and therefore the transfer moment is not necessary an 'internal moment' of the modeling dynamics which contributes directly to the elaboration of the model-based understanding. I argue in this sense against the emphasis placed on the transfer moment as the definitory for the model-based understanding.

In the next chapter, I will try to advance and tackle in a more detailed way some of the ingredients that are to be considered as necessary means in order to sharpen and detail my type of approach on scientific explanation through models.

Chapter 4

Further means to implement the approach on explanatory models

As it came out from the previous chapters, explanation and understanding are to be conceived in a local way according to the modeling processes involved. But what do we pertain to when we talk about locality? I'll try to pinpoint and clarify the ways in which locality is to be conceived and the different meanings in which we encounter it. I will try to distinguish as well as I can these meanings and search for some other more productive for our inquiry.

As I remarked in the previous parts of my thesis, the notion of locality was invoked in a rather informal way, and it appeared in pair of opposites side by side with the term global. Some meanings of the concept are to be understood therefore in opposition to the global notion. To review some of the meanings of global-local that I have already taken into consideration, I should mention the following: the most common refers to the distinction between the appeal to features of larger entities of scientific knowledge – theories, fields etc, even whole corpuses of knowledge[67] bearing the main determination for scientific explanation and understanding versus the invocation of more delimited aspects specific to the situation under investigation. This is the sense that, to put it briefly, Kitcher and Friedman had in mind.

Another meaning of global which can be identified from its use by scientists points to the appeal to most general principles and theoretical ideas – as in the way scientists use in their talk saying that some results were obtained on global considerations. But it could not be this sense that is to be referred to in order to make a difference between models and theories in our case, since models could be built on such general considerations and could also incorporate them.

Another sense implicitly present in the above-mentioned one could be spelled out through the characterization of global as being given through a worldview. Salmon explicitly articulated such an interpretation. Local will not involve such views; but construing the distinction this way will restrict the meaning of the global to a too extreme and limited interpretation.

A way of interpreting the distinction local-global as applied to understanding was suggested by Lambert.[68] According to his claims, if we are after an explanation of a state of affairs and we seek to subsume it into an already known or taken for granted theory, we will have a local kind of understanding. In case we are looking to build a new theory to account for a ‘batch of diverse states-of-affairs’, we engage a global kind of understanding. We notice that the distinction does appeal in its core to the difference between one versus many phenomena targeted. The distinction exhibits also a relativization to the inquirers’ intentions, a fact that will probably make it vulnerable to accusations of

---

[67] Unification is such a global feature.

[68] See his paper *On whether an answer to a why question is an explanation if and only if it yields scientific understanding*.

subjectivism. Some would probably accuse the fact that we would like to have more objective criteria to draw the distinction, than the scientists intentions when engaging in doing science.

In as far as the model is concerned we can also apply the above characterizations in their case. Models could take both positions: the one in which we try to subsume a phenomenon under an already developed model and another in which we try to build a model ranging over diverse states of affairs. This fact will make them participate in providing a local as well as a sort of global understanding. The last claim would be rejected from the perspective of the old accounts, since a common feeling of what the global attribute should describe will not see models as proper entities to be subjected to it. Nevertheless in case an inquirer engages in a model building process he would have to appeal to both local and global understanding but the final product, i.e., the model, will most probably provide a local kind of understanding.

As a more general observation, the above and previous discussions as well as the way the local and global attributes were engaged in the debate, give the impression that locality is to be rendered workable only in a complementary way, i.e., in tandem with the global attribute. I think it would be worthwhile to engage in investigating directly what could be the aspects that were taken more or less explicitly to mark the local character. The global-local complementarity plays its role in subtext but it is worth to try to expose the aspects of scientific inquiry that are tightly connected with this local character alone.

One might consider contextuality as the main mark of locality. The main idea of seeing explanation and understanding as local is to consider them as being determined by contextual factors of inquiry. But some of the senses of contexuality are not welcome since theories and corpuses of knowledge are also subjected to contextual factors. In fact Kitcher's model does accommodate this aspect. The factors at the socio-historical level, those which modulate the scientific enterprise as a whole are not usually the first one intended when pointing to the local character; in this sense, we will say, that they are not the most efficient in that particular situation. The sort of contextuality that is referred to is one that relates to the factors at play in the process of scientific inquiry of a phenomenon in a particular situation. It refers to factors as beliefs, purposes or/and intentions of the inquirer, who engages in the process of scientific inquiry. Hempel and van Fraassen do have this sense in mind when referring to the pragmatics of understanding and explanation. Of course, the factors that influence the macro-level of the scientific enterprise do inform these more local ones but the effective, immediate influence is taken to be from the last ones.

If we are to capture some sense of explanation or understanding at the model level the contextuality of explanation and understanding and the one of models have to meet. Of course, the pragmatics associated with modeling is not the same as the one relevant to explanation; what we are looking for is the intersection zone as the zone of interest. To identify the moments in which their pragmatics coincides will be the objective of the inquiry of explanation and understanding through models.

To recall shortly: a classical reaction to the above claims comes along the line that there is no pragmatics associated with models. Models fall, according to the adherents of the above position, exclusively in a syntactic or semantic register. Two replies are offered to counterbalance these positions. First the building, modification or application contexts are the playground of pragmatics. These processes disclose the explanatory virtues of models. Second, I do side more with the view expressed by R. Giere[69] that the 'scientific practices of representing [through models also] the world are fundamentally pragmatic' and to talk about a syntax or a semantic is a kind of deviant way of putting things. To make further the point, we know that there are some well-articulated positions on explanation available which emphasize the importance of pragmatics. The best known position, van Fraassen's, draws on an interrogative approach on explanation. The means used have to be placed in the context of the modeling processes and I will discuss the subject in a subsequent section.

The attribute local in case of explanation and a sort of understanding could be interpretable also as related to the causal networking revealed through the explanatory process. We have in this case a clear modality to spell out the locality of the explanation and understanding. Salmon, as already presented, is the main advocate of this approach. By opening and revealing the causal mechanisms at work in a particular situation the locality of the explanation is implicitly revealed. Furthermore, Railton brings the idea of explanation as causality to its ultimate consequences in his account on explanation, fixing implicitly the idea of locality.

But there are first-sight drawbacks from several sides. One such drawback is that causality could be invoked in some explanations in a global register too: the sense of global pertaining to scenarios that have a wide range of application. There is room for causal networking in explanatory attempts that promote even worldviews and enhance the global understanding.

The other drawback comes through a straightforward reading of the request for causal information. Railton is the author credited for articulating best this kind of reading. Locality is captured in his view as the completely detailed network of the causal influences at work in the phenomena targeted by the explanation. The critique fired in this case draws on a well-known problematic point: the relevance of the causes. For Strevens the causal approach does not impose any constrains to counterbalance the drive for causal accuracy. For Batterman the local accounts, as Hempel's, or the causal-mechanistic one demand "detail and precision" given through the specification of the initial conditions or of the "nodes and links" of the (causal) mechanism disclosed. This demand induces 'explanatory noise', i.e., too much unnecessary explanatory detailing. The strategy that offers in the framework of a causalist approach to deal with this problem is one that requires an imposition of additional conditions in order to get rid of the unwanted details. Strevens account goes this way, but in the end he draws on the unification feature to induce the necessary constraints.

[69] He expressed his view explicitly in his presentation at PSA meeting 2002.

Batterman offers another kind of proposal[70] that provides an option for the clarification of the locality issue. By analyzing the asymptotic kind of reasoning in physics, Batterman comes to the conclusion of the locality of understanding taken in the following sense:[71] "one achieves explanation and insight by displaying the mathematical emergent asymptotic structures." Even in case of similar problems (as in his examples of diffusion of underground water or the spreading of the waves of perturbations in a liquid) these structures are different. Each problem requires "detail local consideration of the physics involved." The physics of the situation is "singular" and "plays a crucial role in focusing these considerations."

On one side, Batterman's position is a more comprising one, since the singularity of the physics at work in a situation could be expressed in many forms, the causal format being one among others. Nevertheless a precise explication of what it should be understood through the singularity of the physics of a situation is still lacking. Moreover, the physical interpretation of some mathematical structures draw directly on the universality of the physics involved. It is_therefore an immediate need to articulate what does the claim that physics is singular pertain to.

Furthermore, we can see an additional limitation of Batterman's position. It comes from the fact that the account targets a specific form of explanation, i.e., the one given through asymptotic type of reasoning, in which mathematical structures play a central role. The extension of his ideas to other fields of scientific explanation that do not involve mathematical structures is not unproblematic. Nevertheless the morals that Batterman draws from his account regarding the critiques of the traditional accounts (especially the causalist one) prove to be general and effective. But the implementation in a more general register of the main idea of the approach is less plausible.

To summarize shortly the above ideas: as the critique of both Strevens and Batterman showed, unrestrained causality is not a workable solution for a general approach on explanation. In general explanation involves ways of blocking the causal detailing which otherwise will induce explanatory noise. From the straight local perspective the direct reading equates finer casual detailing and the total causal networking gets the complete local story of the phenomenon. But this characterization is not the one relevant for explanatory purposes. Locality in terms of causal networking can at most be a partial explication of explanatory locality. Additional constraints are required in case we are to hold on to a causalist explicatory element as Strevens does. For him the explanatory relevant causal network is ultimately determined by an unificationist constraint, as exposed in his kairetic account on explanation. For Batterman who abandons the causalist strategy the relevance is delivered through the emergent mathematical structures.

Returning to our plea, I think that a better perspective on the locality issue could be gained in the modeling framework. In a model-oriented approach the issue of locality is better rendered through the locality of the modeling process. We can avoid this way the pitfalls of some approaches as the ones

[70] In his book *Devil in Details*, but also in some other papers such as *Towards a 'Modern' (=Victorian) Attitude towards Scientific Explanation*.

[71] See *Towards a 'modern' (=Victorian) attitude towards scientific explanation*.

mentioned above. The modeling perspective offers us a more general modality to capture the locality of explanation. It is general in the sense that it could account for more specific forms (the causal one being among them) of pinning down locality. The pragmatic approach on locality can also find a natural place under a modeling view, as pragmatics is embedded into the model entities. An extreme interpretation of locality and a possible divergence from the explanatory one as in case of the causal approach is avoided by the fact that the explanatory relevance is given through the model; the locality of the explanation is given through the locality of the modeling process. A further more refined explication of the locality of explanation is to be articulated within the frame of the locality as expressed through models.

Local unification

There is a feature so central to the explanation process that it has to be reconsidered under the view I'm pleading for. Unification was seen exclusively to belong to the global register, but in considering explanation at a local level there is unification and coherence to be claimed in this case too. Besides, if we are to take the coherence as a mark of understanding we should look for the sorts of unity induced this way, i.e., which are exhibited at local level. Models would play an important role in this sense through their contribution to the clarification of what unification and coherence at the local level could be.

And indeed how could we conceive unification at the local level? A first look will direct us to search for analogies with the unification concepts already developed. It is possible that the sorts of unification we encounter at the theory's level could be also realized in models. In fact there are features that not only theories exhibit; some specific models – as the highly theoretical ones will exhibit them too. Regarding the characterization of unification realized by theories, as I already mentioned in my previous chapters, Friedman's reductive relations could be realized in some types of theoretical models. The same could well hold in the case of Kitcher's (but also Schurz') concepts of unification. Possible refrains to the transfer of these notions in models' register could invoke a kind of 'unrestrictedness' regarding the facts to be integrated through a global unification. In the case of models there seem to be some important additional constraints; this is true especially for unifications via larger knowledge entities (as the corpuses of knowledge), since theories will also exhibit some constraints for what specific facts they provide understanding. But at least the central ideas of the above mentioned accounts could make for a starting point or a comparative guidance to unification through models and in general nothing so strong can be found that will restrict any such application. Consequentially, it seems that applications of these notions to models are in need of specification of more constraints coming from models.

Morrison argues in her book *Unifying Scientific Theories: Physical Concepts and Mathematical*, for many existing strategies of unification. Though she decouples the topic of explanation from the one of unification, her conclusions on unification are relevant regardless of her position towards explanation. In her conclusions to the book, Morrison emphasizes that there is no unique mode to characterize unification. But, on the other side, in each example she puts forward, there is a general and abstract mathematical frame that is mainly invoked to do the job: “In each of the examples I have discussed, a general and rather abstract mathematical framework has played an essential role in the unification of disparate phenomena.”[72] Even in the case of Darwin's theory, she discusses with the intention to strengthen her main point with a theory that lacks explicit abstract mathematization, the unificatory potential behind the natural selection draws on the quantitative structure of Malthus. In the case of models, to reduce the unification aspect to the presence of a mathematical structure is, of course, a restrictive and specific view on the issue.

But, there is another more suited way of viewing the unification realized through models. Morrison leaves room in her conclusive remarks for other sorts of unity as for example unity in practice – which she acknowledges as being different from what she discussed in the book: “there are other forms of unity that may exist to greater or lesser degrees at the general level of scientific practice.”[73] From the perspective of our investigation, the suspicion rises that the unity realized through models falls mainly in some of these other registers. But what should this mean more precisely and how to approach such other forms of unity is left unspecified by Morrison. I will try to make some suggestions on this point later in this section.

One could refer directly to the explanatory relations as responsible for realizing such local unity. V. Gijsbers follows this track in his contribution to the 2005 *Conference on Philosophical Perspective on Scientific Understanding* (ideas that will be later restated in his paper)[74]. Explanation is construed as a determination relation between facts; F the *explanandum* fact and a set of facts G making out the *explanans*, both F and G being selected from specific contrast classes. An explanation would exhibit two main steps: one involving spelling out what G contains and another showing how G determines F. Unification is realized locally due to a connection between F and the set G. Local connectedness of F is enough to explain it, though it can be that the explanation is disunifying with regard to global unification, i.e., it lowers the overall unification. Usually we reject disunifying explanation because of some the statements involved are in conflict with other parts of knowledge; “but *if* – for whatever reasons – we are willing to accept the premises of a disunifying explanation, it can function perfectly well as an explanation.” Global unification seems to be a hallmark for good explanation for two major reasons, according to Gijsbers’ because they generate many other explanations and because their premises are more likely to be believed. But “neither of these is a necessary condition on explanations per se.”[75]

Though his argumentation seems plausible in what regards the critique of global unification and also in making the point for a local one, his articulation of the local explanatory relation remains in general unsatisfactory. The main cause of this is the fact that the relation of determination is left totally

72 In *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*, p. 233.
73 *Idem*, p. 235.
74 In *Why unification is neither necessary nor sufficient for explanation*
75 Gijsbers presentations-slides at the Conference On Scientific Understanding Amsterdam 2005.

unspecified. I would see as a necessary further step in the frame of his endeavor an appeal to a modeling view. Bringing a modeling view to bear upon Gijsbers' ideas will provide the missing articulation or at least create the proper frame for this investigation. As some possible suggestions: the *explanada*, i.e., the set of facts G would have to be spelled out in the modeling terms as for example the representations of these facts would be part of the ones that make out the model. The local coherence obtained through the local connection would be reflected in the coherence realized by the model.[76]

In order to get a grip on the unification aspects at the model level, I think, a good strategy could be initiated by placing unification in a dynamic context of scientific inquiry. In the camp of practice-oriented approach to modeling, Boumans offered a description of the practice of model building from which some suggestions can be drawn for the issue of unification. I will present briefly Boumans' account and further extract the morals for my inquiry. In his contribution to the volume 'Models as Mediators'[77] Boumans compares model building with "baking a cake without recipe". The main thesis claims that model building is better viewed as a process of integration of ingredients of various natures. Models have to meet implicit criteria of adequacy and in order to do this they have "to integrate enough items". The items are very different in nature: theoretical notions, metaphors, analogies, mathematical concepts and techniques, policy views etc. His claims are exemplified through mathematical models from economics: Kalecki's business-cycle model, Frisch's Rocking Horse Model and Lucas' equilibrium model of business cycle. In the cases discussed the criteria of adequacy are made out of a selection from the following (not complete) list of requirements: a solution to theoretical problems; an explanation of empirical phenomena; an indication of the economic policy; a provider of a mathematical conception.

The paper addresses also directly the issue of discovery and justification in the context of modeling practice. The context of discovery is one of a successful integration of the items in the process of model building. As a by-product thesis, he is stating the claim of the inadequacy of viewing discovery and justification in a separate way. The idea of the built-in justification is advanced in order to clarify the new place of justification. I'll address shortly this issue at the end of this section.

The main point that I would like to make by drawing on Boumans' paper is that if we are to see the unificational practice at the model level, we have a good starting point in viewing unification in models as integration of various ingredients under a modeling purpose. What can such a view offer us? The most important thing is that it exhibits the advantage of centering the investigation in the process of scientific inquiry.

One could react that it is nevertheless a particular process, i.e., the one of model building, and that it does not capture the variety of senses exhibited by unification. In reply it could be said that this approach is not meant to be an absolute one in the sense that it targets all sorts of unification existing at the model level. It is rather the special importance of the process of model building in placing models in

[76] Though not totally explicit there is a tendency in Gijsbers' account to see unification realized through the coherence of beliefs. To secure the objectivity of such an account we will have to add some more constraints.

[77] *Built-in Justification*.

the frame of scientific inquiry that qualifies it as the most interesting for such a job. Other processes such as model modification, calibration or model application could be seen as variants of the model building process. A description of such a process will of course have to take primarily into account the constraints under which the building takes place.

But one could ask what is the difference from the descriptions of unification as were previously given - i.e., would it not boil down in the end to some similar description as the ones we know for unification? The main difference is to be found in the fact that in these approaches we account for unification as a final product, while integration makes us look at unification as an activity. This allows a better accommodation for the elements of pragmatics that play such an important role in the modeling process. So, for example, the purpose for which a model is built can be made explicit under this approach.

A short adjacent remark regarding the built-in characterization of pragmatics as advanced by Boumans: there is a sense in which pragmatics is built-in in need to be explicitly disclosed. It is the pragmatics at play in the in the process of model building that gets frozen in the final product – the model as it is presented in its general lines. This has to be further clarified in greater detail.

Another reticence on using this approach comes from the fact that for Boumans the explanatory requirements are only some among others that are to be satisfied. This qualification is justified in general since not all models do have an explanatory objective. But of course the reticence is not justified if we restrict ourselves to the explanatory models, i.e., the ones that impose explanatory requirements. A pan-explanatory position would have to establish the existence of implicit explanatory virtues besides the existence of explanatory requirements or take such requirements as a constant of any modeling process. It seems to me that a restrictive approach is advisable for the beginning, and this will induce the fact that explanatory unification is only one brand of unification in the modeling context.

There is a clear limitation in Boumans' proposal due to the explicit reference to the mathematical structures that play the central role in his examples. It parallels Morrison's claims from her analysis of the unification at the theory's level. For Boumans integration takes place 'by transforming the ingredients into a mathematical form to merge them into one framework'. He illustrates this claim through the mathematical entities that make out the economic models. The immediate question that arises is to what extension can we make this claim for other types of models? The claim appears to be valid in general; the problem would be to find some equivalent for the mathematical structures. There is a hope in the subtext when pointing to this issue. It lies in the fact that the type of the models will disclose the sort of structures that imparts the necessary unity.

The methodological question is related to the modality to disclose such structures. The direct and simple way to do it is to appeal to a robust view, as the semantic view on theories. The unificatory structures are the ones that are referred to in this view on theories. I think one does not have to fear in this case that the unity in the models is derived from the one in the theories; there is enough room to claim the autonomy for such a unity through models.

There is another way to render the unity realized in the models, in a pragmatics-oriented register. Boumans' unity is of a theoretical kind; a sort of practice-centered unity that could be illuminated by an appeal to the purposes for which models are built. But how does such a unity through purpose relate to the explanation or understanding we get through models? It depends on how we relate the explanatory virtues to the purposes that models are built for, and that's what I'll try to address after a short parenthesis.

A parenthesis: the discovery-justification distinction and the built-in aspect

The objective of Boumans' paper was to argue also for the inadequacy of the traditional separation between the context of discovery and the one of justification. The separation was already challenged and dismissed through previous critiques. Hoyningen-Huene[78] exposed the different conflated modalities of reading the distinction and the basic assumptions behind the distinction. The inadequacy of distinction is therefore not an issue; the novelty comes from the specific application in case of models. In Boumans' view one could say that the context of justification is dissolved in the one of discovery. The overlapping of the processes would have been allowed also by the classical positions that promoted the distinction. What is most probably meant by Boumans in drawing on this distinction is better made explicit, I think, in the way Hoyningen-Huene viewed the core of the distinction, i.e., as being between two perspectives: one referring to descriptive aspects and another related to the normative-evaluative aspects. The normative-evaluative aspects are incorporated into the model in the way they guide the constructive process; through the *a priori* set criteria the descriptive ingredients are adjusted and accommodated to fit into the frame of modeling.

There is a further perspective to expand to new alternative proposals that recuperates on a less 'classical' position a separation between contexts. So for example, Echeverria[79] proposed the following contexts to be taken into consideration instead of the classical distinction: the contexts of education, of innovation, of evaluation and of application. The framework in which he advances these distinctions is in agreement with the broad assumptions under which the inquiry of explanatory models could and most probably should be undertaken. The framework[80] comprises such theses as: the one that philosophy of science should analyze real scientific processes, that scientific process involves many other processes besides discovery and justification, that a dynamical approach to science should be preferred to a static one, and scientific activities and not scientific knowledge should become the primary concern of the philosophers.

Now we could ask about the perspectives that are induced for our inquiry by adopting such a new partition. I will not evaluate the fruitfulness or quality of this distinction *per se* but try to sketch the consequences in case we will take it as a working hypothesis. The new typology of contexts will

78 In *On Varieties of the Distinction between the Context of Discovery and the Context of Justification*.
79 In *The Four Contexts of Scientific Activity*.
80 *Idem*, p. 156.

reshape the ways of approaching explanation and understanding. So, in the context of education it seems to be possible to account straightforwardly in a detailed and articulated manner for the explanatory and understanding acts. The particularity of the processes in the educational context simplifies the problem, while the known means, as for example the ones provided by interrogative logic, become more effective.

The context of application is meant to characterize the applied zone of scientific activity. This will focus our attention on how models are built especially in engineering fields. The criteria of successful modeling will have to consider the functional dimension of the models as playing their role in the applications. As explanation and understanding involve more theoretical activities some people may have doubts in invoking them in such cases. Of course no one will totally deny the existence of explanatory acts in applied science, but they will be regarded as secondary or rather subordinated goals to the other.

Innovation and evaluation seem to restate in a different way the classical distinction between discovery and justification. But what should make it avoid the pitfalls of the old distinction is the specific scientific activities in which they manifest. Escheverria sketches some suggestions of how these two contexts can be read out and acquire concrete shape in scientific activities. But the important point I want to emphasize is that such a differentiation of contexts would also induce in a way a typology of major purposes of scientific modeling. It will be further in need of enrichment with some finer distinctions. In this way concerns on mixed goals such as for example building a model in the educational context that could play a sort of an evaluative role could be better qualified.

The purposes of models

It was claimed that models are representations that should be characterized as having attached a purpose. In spite of the communality of the claim nobody tried to discuss how it would accommodate the explanation and understanding processes in this frame. Further on I will try to tackle the idea of situating explanation and understanding under the view of models as entities that incorporate a purpose in their constitution. It is an obvious fact that purposes range over a broad spectrum – they comprise a variety of intentions to use the models. A first issue that arises is related to the quest whether we should take explanation and understanding as simple purposes among the others, i.e., on the same footing as the others or take them in a totally different regime. The quest arises due to the huge importance that these processes play in the economy of scientific inquiry. There are some bold cues that will make us incline to situate the two on another level than the one of other purposes.

The first most important cue is the fact that explanation and understanding purposes underlie in a fundamental way the entire scientific effort. What drives the scientists in their inquiry activity is exactly the fact that they want to explain and understand the phenomena under scrutiny. Another reason that can be pointed to lies in the way the usual purposes are realized. Simple identifiable tasks and objectives, ones that can be clearly articulated are usually indicative for the usual purposes other than explanation or understanding. In the case of explanation and understanding such a claim cannot be usually made in a sufficient explicit way.

For example a concrete purpose for building a predictive model is to enable the generation of reliable predictions for the behavior of the target system. The objective is clearly identifiable and can be spelled out without difficulty, giving also criteria that mark the accomplishing of the task. In case of explanation it is not clear – in spite of the rich debate around the subject – what kind of task we should perform or what sort of performance the model should exhibit. Criteria for signaling the fulfilling of the task are also quite unclear.[81] Sometimes various more concrete and better identifiable tasks accompany in specific situations the explanatory accomplishment through models. But neither of these makes for a constant propriety in order to characterize in a necessary and sufficient way the explanatory purpose. Besides, it is not clear when the goal is indeed achieved; this is reflected also through the lack of sufficiently general criteria to identify it.

A first problem that we can raise is by asking in what kind of relations do the other purposes stay to the explanatory goal (that of understanding being included)? Some were intimately associated with a partial realization of explanatory goals. How this partial dependence is to be spelled out is given by the particularity of the circumstances in which the realization takes place. We could say that it happened that criteria were transferred from other goals (such as prediction or unification). We could claim also in a good sense that the purposes of explanation and unification are elusive in comparison with other scientific activities.

The more fundamental role that explanatory inferences play is emphasized by P Lipton when making the case for the inference to the best explanation. He points to the fact that even in such cases where explaining is not our primary purpose, an 'explanatory detour' is involved. In his words: "the explanatory detour seems to be one of the sources of the great predictive and manipulative successes of many areas of science."[82] How does such a detour enhance the realization of a particular purpose is a question that has to be addressed not only in this inquiry context but more general on any further attempt on explanation problem.

One point to start is to notice that there are some purposes that have a more intimate contact with the two fundamental processes. One is the descriptive purpose. Explanation seems to involve under any interpretation a descriptive act. As van Fraassen puts it, we get in fact a description every time we make a request for an explanation. But it appears that description makes for a necessary condition but not for a sufficient one too. We cannot account for explanation by accounting only for the descriptive activity. On the other side if our scientific activity is explanation-driven in the most fundamental way, scientific descriptive performance should exhibit explanatory elements too. It was claimed on many occasions that the distinction description-explanation cannot be held clear-cut. This is

81 Sometimes we encounter such justification as the one of theoretical physicists who invoke the beauty or elegance of a formula or theory as a reason for being explanatory.

82 P. Lipton, *Inference to the best Explanation*.

not an issue anymore. The problem is under what circumstances we can engage in using such a distinction. One point that I want to stress in this sense is related to the fact that there is more explanatory virtue than it was usually attributed in the descriptive activity especially when it takes the form of a model. A descriptive model encapsulates explanatory virtues, comprising lines for developing new ones. It provides also a frame for performing explanatory activities.

I will discuss in greater detail the above claim with reference to an example which seems to be less common for the thematic of models and their explanatory qualities: an organism model. For this purpose I am taking the worm C-elegans as the concrete example of such an organism model. It is one of the models widely used in biological studies in actual research. One might think that such a type of model would not exhibit any explanatory virtues. Even such more recent approaches as Ankeny's[83] tend to strengthen the above claim. I want first to emphasize the interesting features that such kind of models (i.e., material ones) exhibit under a more general view. The focus in the discussion of models in the philosophy of science was mainly on mathematical sort of models. Recently many authors have pointed to the importance of taking into consideration other types of models too. Sterrett[84] is one of the main advocates of this point backed by her investigation on other sorts of models than mathematical ones. She calls such models 'piece of the world' kind of models and points to the fact that there are often quite 'formal methods of showing that one concrete thing models another'. This fact shows that there are non-trivial aspects to be recovered from such scientific episodes of modeling from which the explanatory aspects are not to be excluded.

C-elegans is a tiny worm, free-living nematode which is nowadays one of the most important model organisms in the molecular biology research. The importance is emphasized by the fact that it was the first among the model-organisms chosen in the Genome Sequencing Project to be entirely mapped. In comparison with other model organisms, it is a newcomer in the field since it was introduced by Sidney Brenner relative recently, in the '60. The reasons for choosing it as a model organism were also contextually determined besides the usual consideration for any model organism. The general consideration includes such items as experimental tractability, similarity with other (higher) organisms, availability of background information etc. In case of C-elegans a main contextual reason for choosing it was given by the necessity to explore the new (at that time, in the '60s) area of research of neurobiology. The main goal was to connect genetics with behavior at the level of an organism as Brenner himself presents it "the hope was to dissect the genetic specification of behavior" in the same way as it was done with biosynthetic pathways in bacteria. The already used model organisms as bacteria were limited in this sense; so a sufficient simple organism was required which should also have a sufficiently complex neuronal system. Though at the end of the first period of the research it became clear that causal information of neurobiology was not obtained as it was hoped, the emphasis moved to description, laying this way the foundations for the correlation between genes and behavior.

Ankeny documented in detail the research on C-elegans from its inception as a model organism

---

83 As articulated in *Fashioning Descriptive Models in Biology: of Worms and Wiring Diagrams*.
84 In *Kinds of Models* and *Models of Machines* and *Models of Phenomena*.

in biological research and advances also some philosophical and methodological conclusions related to this sort of research. Her argumentation intends to emphasize the importance of the concept of descriptive model for understanding the scientific practice in this case. She views the history of the research on C-elegans as an episode of scientific activity that develops in a pre-explanatory register. The main point she wants to make is that there must be a pre-explanatory phase in which a descriptive model of the experimental organism is developed. It produces a sort of standard to which further explanations make reference when they are developed.

A descriptive model is conceived to be an abstract and idealized entity. The concept of descriptive model can be seen as a special instance of the process of extracting phenomena from data; the last ones are idiosyncratic to a particular experimental context, while the phenomena are characterized by stable repeatable features. Ankeny illustrates these points through the wiring diagram, i.e., the complete structure and connectivity of the nervous system of C-elegans.

Approximation, simplification, and idealization were used in many places in the process of elaborating the wiring diagram. Some of such steps are further represented. The wiring diagram is in fact a mosaic of nervous systems of four individual worms. The individuals were used to produce series of sections that proved to overlap generally, and also to exhibit some variation, which had to be eliminated. The reading of the micrographs was based on different interpretative assumptions. The connectivity data was exposed through classes of neurons though there was a low-variability with respect to the number of synapses. Other difficulties included the prediction of neuronal function from neuro-anatomy or lack of functional type of particular synapses.

The point I want to make and throw this way another light on her conclusions is that explanatory virtues are not to be excluded even from such episodes in which descriptive models are elaborated. I want to amend and further qualify her claims in the light of the fact that the explanatory practices underlie in a fundamental way the scientific practice. Her claims are justified only under specific qualification, i.e., in case we are interested in fixing a particular explanatory episode and ignore other levels of explanation. Working under this qualification we should keep as proviso the fact that any modeling level could be taken as a descriptive one relative to the levels above them and as an explanatory one in relation to the lower levels. The case for the explanatory virtues associated with the wiring diagram can be made from different points of view. In principle we can make use of any of the well-known explanatory approaches on the wiring diagram example to reveal the explanatory virtues at work. So the DN model will give us the subsumption under the lawful generalizations that are embedded into the descriptive model. The model itself could be identified with such a set of lawful generalizations even if at a lower, more phenomenological level. Kitcher's explanatory patterns are applied during the construction of the model and are also incorporated into it. On one side we notice the more general patterns from the field of the neuroscience that apply in this case. As we move to more recent and generous views the fit seems to work further (or at least does not seem to be problematic). M Morrison's claims on explanation through models, which I discussed in a previous chapter are applicable in this case. The LOOP account as proposed by Hartmann and Frigg seems also to be adequate for disclosing the explanation in this situation. In the case of a particular worm, the wiring diagram gives us a picture in which a correspondent OIM for an occurrence in the system can be obtained. In general the explanatory effects are to be seen in the way the neuronal structure and behavior of a particular system – in our case an individual worm – is rendered intelligible through the wiring diagram.

A substantial proof comes from such a well-articulated view on science and scientific explanation as the structuralist view. The approach on explanation advanced in the frame of the structuralist conception integrates nicely such results. Bartelborth's or Sneed's view on explanation will attribute explanatory qualities to such descriptive models. Explanation through embedding will render the wiring diagram as an explanatory model. In a more general way the fundamental explanatory nature of scientific activity is rightly exposed under the structuralist view. The view gains also more weight by being backed through results from cognitive science studies related especially to the topic of conceptual knowledge.

The fact that such model organisms as C-elegans are exhibiting explanatory virtues can be best seen in the frame of articulated accounts on explanation that take as C-elegans as their working example. Such is Schaffner's account for explanation in biosciences and medicine; I'll further discuss his account mainly with the aim to extract further clues on the explanatory roles models can play.

K. Schaffner picks directly on the scientific explanation related to the research on C-elegans.[85] He applies his model of explanation in bioscience and medicine to this particular case. He is focusing on two types of successful explanatory episodes from the biological research on C-elegans. These kinds of explanations reflect at the same time two options for different research programs. One takes as its explanatory goal the genetic explanation of the social behavior (following in this sense in Brenner's legacy); the other one ignores *explananda* in terms of genes in favor of a complex neurophysiological account of the behavior. The first explanation proposed in a study of De Bono and Bargmann identifies a gene mutation that causes the social behavior of the worm: i.e., it switches between a solitary and a social behavior of the worm. The other type of explanation, due to Ferree and Lockery, is given through a neural network model for the chemotaxis control circuit in C-elegans. The simulations of the solutions of the equations proved to be in quite a good correspondence with the experimental trajectories. This last type of explanation resembles rather explanations that are characteristic for biophysics with no appeal to genes or molecular biology.

Schaffner brings together both types of explanation under his account of scientific explanation. His approach conceives such an explanation to have two elements a focal and a field one, (called as such the 2F approach); he is also trying this way to strike a compromise between what he describes as the narrow models of explanation, as for example the DN model and the broader ones, in which explanation is built in larger contexts of knowledge such as in Kuhn's paradigms, Kitcher's practices or van Fraassen's interrogative model. Making use of the two types mentioned earlier, the genetic type of explanation, though it is very specific in providing explanation of some behavior of the worms, is

[85] In *Behavior at the Organismal and Molecular Levels: The Case of C- elegans.*

backed by some general background assumptions and allows us to focus on some more definite questions. In the context of the answers that give general reasons for the research undertaken, we can pick out a focus and ask about the nature of the explanation in a narrow sense. It is here, as a last step according to Schaffner that we can appeal to laws, mechanisms or models or refer to the nature of the inference as it is the case in classical accounts on explanation.

In Schaffner account each of the two elements, the focus and the field, have two components. The two main focal components are a model system and a causal component. The two field-oriented components are reflected through a comparative-evaluational-inductive aspect and the ideal explanatory text background. As we can notice models play a central role in his approach. The model system component refers to an idealized system that contains also general assumptions about the system under study; explanation is provided through its instantiation in particular cases. Schaffner claims to have in mind the kind of semantic approach when referring to the model system.

For the second focal component, i.e., the causal component may be characterized according to Schaffner as a kind of necessity and sufficiency “in the circumstances”. In case of the deterministic causation he thinks that deductive logic suffices to bind together the model system and its instantiations, while in case of a probabilistic causation some Bayesian type of support is required.

The first field component, the comparative evaluative inductive one, is meant to recuperate the implicit evaluational component, which Schaffner finds to be present in all accounts on explanation (e.g., truth for Hempel, evidential support for van Fraassen). It is best construed in his opinion as ‘doubly comparative’ in the sense that: “for a given explanation one compares (often implicitly) an *explanadum* with its contrast class in terms of the support provided”. The second field component, the ideal explanatory text background reiterates Railton's main idea. Applied to the two explanations on C-elegans, only a small part of the ideal explanatory text is disclosed through them.

Now regarding the role of models in explanation, Schaffner's account makes explicit appeal to models and gives them a central role. To stress this claim we see that in the context of the application of his account to the two types of explanation on C-elegans, Schaffner points to the main distinctions between the explanations as lying in the different model systems appealed to. So in fact the very identity of the explanation is expressed rather in the model component rather than in the other components of his explanatory setting.

We could have in principle different field components too – probably if considering explanations from very different areas of science. As such someone could invoke models as difference makers and specific explanation-carriers, but only after we settle the other components by delimiting eventually more precisely the area of the research. In case of Ankeny’s claims the building of wiring diagram will make for such a different explanation due to the different model-system appealed. Though Schaffner makes reference in his paper to C-elegans at different levels of data gathering and analysis, he does not advance the claim for the existence of a pure descriptive pre-explanatory stage. Rather he goes on to stress the different levels of modeling of the circuits. The ideal system seems to play the kind of the role an explanatory schema played in classical accounts. As it stands there is an obvious need to connect it to the more comprising scientific field and focus structure in a tighter way than by simply associating it with the other components. I think that a solution would be to draw on a semantic view notion of models. This way we could also recuperate better the explanatory virtues of models under this setting especially in the frame of the well-articulated structuralist account on explanation as an embedding.

Now taking into account Schaffner's account *per se*, the main general critical point that can be made regarding Schaffner's approach on explanation is, I think, that his view is too dispersed. The ingredients are spelled out separately in a sufficiently detailed way, still the drawback comes from putting them together. The account gives us rather a sense of listing of the components without providing a synthesis of them, a way in which they are connected and integrated under a schema. The listing lacks a justification of why these are the elements needed and not any other or why the list should be complete. His account could be taken as a wish list of what elements should be included in an approach on explanation. From another perspective and by reading it in a key that takes models as explanatory bearers we can see Schaffner's account as signaling the need to enrich a too restricted view on models. That is in a good sense the main message the account can pass on to us about the relation between explanation and model and what is the need to situate the models in a larger enriched context of inquiry. And what I mean by this is following:

Such as it is conceived, the account does not make explicit room for any pragmatics. From amongst all components, the one that is most open to pragmatic influences is the first field component, the evaluation inferential component. The needed enrichment would have to induce the pragmatization of the account by referring to this component. Purposes of the modeling processes for example would have to be reflected primarily by this component.

A more synthetic view for Schaffner's account could be reached through a fully model-oriented approach on explanation. This will resettle the elements spelled out separately in a unificatory modeling framework. As a direct consequence this step will require a more generous notion of model than the one Schaffner adopts which is conceived as a set of generalizations. As I already touched in some other places, the causal element can be seen as a characteristic of special types of models – the ones that present explanation through causal mechanisms. But Schaffner wanted to make room for other types of explanation too; he points to the equilibrium types of explanation in biology and includes in one variant of his approach[86] a unificatory component.

In order to realize a synthetic view the conception will need an idea that should function as a kind of binding cement. It should be a general enough idea in order to be implemented for any sort of model and it should capture the real mechanisms at work when we engage in the process of model building or model modification. Such an idea should foster an approach that should directly address and illuminate the scientific inquiry as a central process of reference. I think that such a frame could be found in the interrogative idea. I'll turn further to discuss in more detail this proposal and argue for its appropriateness and the advantages of taking it seriously into consideration.

---

86 See his book *Discovery and Explanation in Biology and Medicine*.

Before going further to discuss the interrogative component, I would like to end these remarks on purposes by bringing into discussion a sort of purpose which was recalled by some authors as being a primary purpose when using models. It is the purpose of learning. Giere is one of the best-known authors who emphasize this purpose as the main purpose of models. In his words "models are being used for the general purpose of simply learning what something is like". Mary Morgan does also stress in several places on the importance of learning from models. She offers us a more detailed analysis of how this learning is achieved.[87] As stipulated in the conclusions of the paper, there are two distinctive moments when we learn from models: by building them and by using them. Learning by building models involves finding out what fits together and could represent some aspects of a theory or of the world. She exemplifies her claims by a reference to an episode from the history of economics regarding Irving Fisher's economic models. The building process involves constructing analogies and making choices that are guided by these analogies. By exploiting the neutral analogies and trying to neutralize the negative one, the modeler engages in making different choices from which he learns. In the case of learning by using models our learning is dependent on how much structure from the model is transferable to the real system. Through manipulation of the representations in the model, the modeler acquires knowledge of certain features of the real system and the way he can intervene to control them. It is in such a sense for Mary Morgan that "the power of representation is intimately connected with the means of learning".

By comparison to other authors, Mary Morgan offers us a full and detailed exemplified case of how the learning process takes place in the modeling activity. Indeed learning could be claimed as a main purpose of modeling but at a cost. The less adequate part of viewing learning as main purpose is the fact that it stands for a too general and universal purpose. Any other means that we use to interact with reality are engaging us also in a learning experience. In order to approach in an effective way the explanatory virtues of modeling we need to take into consideration the specific purposes. The task of rendering this specificity brings us back to the more particular activities unfolded in the context of the scientific inquiry. In fact, Giere seems to be aware of this difference when he mentioned besides the general purpose of learning, also the specific purposes that models could have. This allows us to make sense also of the form Giere sees as making explicit the representing activity: "S uses M to represent W for purpose P". It shows that P can stand for specific purposes not only for the purpose of learning.

But from a more general perspective, for the naturalists who take learning as a fundamental purpose, as Trout's remarks on the topic also suggested, learning seems to make for an attractive item on their agenda. Learning as a cognitive process appears more approachable through scientific means than any other activities considered. Besides, Giere has in mind also the other specific types of purposes when he pleads for the naturalization of intentionality. Naturalism seems to provide a general framework in which the investigation of purposes could be pursued. But despite the calls of the naturalists, given the actual research and the results in the field of explanation problem, the prospects for a consistent and widely accepted working agenda in this direction are still elusive.

87 In *Learning from Models*, her contribution to *Models as Mediators*.

Interrogativism – a necessary ingredient

One of the main voices who pleaded in favor of considering the interrogative approach central to the problem of explanation is Sintonen. Over the last two decades he advocated constantly for the interrogative position. His position raises and gains even more weight for the kind of approach on explanation problem I advocate for, if viewed against the background of Hintikka's account on scientific inquiry (on which Sintonen actually draws). I'll restate here some of his points amending them according to the direction of the agenda I'm pursuing. I'll discuss also this way the reasons that speak for engaging interrogative means to render properly the explanatory quantities of the models.

In a 1999 paper replying to Koertage's remark on the problem of explanations, Sintonen underlines the main advantage of an appeal to the interrogative approach, and thinks that it provides the best overall frame in which all the other accounts can be compared. Of course this fact does not directly speak for its immediate usefulness in considering the specificity of the explanation through models but it points to the fundamental character that such a view exhibits through the fact that it provides a base for any sort of account on explanation. With this last reason in mind we can see one of the main motivations for engaging this view in our inquiry.

Another boldly expressed reason in favor of considering interrogative aspects in our inquiry is the following: since the model-oriented view is meant to cover the gap between the scientific practice and the philosophy of science, it moves in the same direction as interrogativism. Therefore a coupling of the two approaches should enhance the power of the inquiry. As such it seems natural to ask why the interrogative and the modeling approaches had separate evolutions or, in other words, why is it not the case for the adherents of one position to appeal to the means of the other?

One reason for this lies in the fact that the two approaches appeal to totally opposite background attitudes. The attitude favored by the interrogative approach is one that promotes formal means in the search for a solution. In Sintonen's words "the interrogative idea continues the research program of the logical analysis of which Hempel's model was an outcome". The modeling approach claims itself to be the continuator of the reaction against the received view in the philosophy of science and accordingly plays down the role of formal means in philosophical investigation.

The rejection of the working agenda of the received view and implicitly of the topic of explanation as a central topic of philosophical investigation was emphasized earlier. As the interrogative approach was invoked under the explanation topic, it did not find a use among the tools of the model-oriented adherents. A brief parenthesis is in place here: not all philosophical views which take models as central for understanding science reject also the appeal to interrogative means. A conception as the semantic view, as articulated by van Fraassen, promotes also one of the boldest interrogative approach on scientific explanation. But of course this is a quite different way of conceiving models than the pragmatic view referred usually under modeling approach. In this case we have quite a regimentation of the modeling approach under a formal frame; while in such a case, as van Fraassen's, we do not have an account that blends the two means, i.e., models and interrogations, together in the inquiry. They are used in separate ways and for different purposes in the economy of his philosophical conception on science.

A common point on which both approaches agree is the important role given to the context of scientific activity. The two approaches agree in the fact that one can say more about scientific explanation if the contextual factors are given a prime role. There is also one comprising idea on which both approaches agree; the one that takes the process of scientific inquiry as the central piece on which the analysis should focus.

Another line in which both approaches meet is the possibility of recuperation of the heuristic for the process of scientific discovery. In Sintonen's words "the interrogative view helps to appreciate the recent recovery of the heuristic and logic of discovery". As I pointed out in various places, the model-oriented view can accommodate these aspects in a natural way.

A point of divergence might arise as to the way the role of theories is seen in the two approaches. The model view as it is especially promoted by the pragmatic view on models, takes a stand against a theory-centric view on science. The interrogative approach does not manifest an attitude of rejection towards theory-centered views. Sintonen is one of the authors who argue explicitly for the need of a well-articulated account on theories in order to make the interrogative approach workable for the explanation problem. For him the main rationale for the request is: "to see how questions emerge from a theoretical background" so that "once the interrogative perspective is embedded in a richer notion of a theory the impetus for expanding and revising the network comes from within, not simply from outward observations and experimental results." This move makes for a naturalization of the explanation problem in the frame of a theory of theories. Under this view there is a potential conflict between the two approaches especially in reference to the model-oriented positions that advocate for the autonomy of models from theory. But the conflict is not a necessary one; we may need a good theory of theories in order to account for the explanatory virtues of theoretical models, but this may not be the case in accounting for phenomenological or other types of models.

On the other side, there is a clear need for more structure when involving the interrogative approach. Since the chances of a general theory of models are small, there is little hope to get the structure from this side. The hope shall come from the rather particular type of models and modeling strategies in which the problem should be settled. In this sense a modeling view can provide the necessary "structure."

Now, there are clear advantages in my view to appeal to interrogative means in order to pursue my inquiry. I already mentioned many of these advantages in the previous chapters; they apply directly to investigation suggesting promising results. I'll touch directly on these advantages in this section. To mention only the two most important ones: one of them is that this view allows the recuperation of the processuality dimension of explanation. The interrogative toolkit allows us to investigate and illuminate such a process, since building an explanation follows a sequence of questions-answers steps. Another such direct advantage can be emphasized in the light of what we discussed in the previous sections in connection with the purposes of models. It appears natural to pin down the explanation purpose by using interrogative means at least partially but also making for a good fixed point for the inquiry, the purpose of an explanatory model is to provide an answer to a why-question. In the frame of a more specific interrogative account on scientific inquiry, as Hintikka's interrogative model of scientific inquiry (the IM), the above idea takes on a clear articulation. By appealing to the interrogative frame we can also achieve a partial disambiguation of the pragmatic concept of purpose of the modeling processes.

There are some direct objections that could be raised against considering the interrogative view for the purpose of investigating explanation through modeling. In general we can expect that the objections reiterate those already raised in the explanation debate against the interrogative approach, so they find a partial response in that context. The main one accuses the fact that the interrogative means are too weak to forge such a topic as scientific explanation; the weakness of the logic of why-questions is especially invoked on this behalf.

In order to reply to these objections one has to point to the essential character of explanation i.e., being an answer to a why-question. The objection could not be pushed anyway in a strong sense since it is based on the received view requirements of capturing explanation only by formal means. The formal articulation of the interrogative means (through the interrogative logic) could not be regarded as a substitute for a formal explication of explanation. As the general accepted critique of received view showed there is a need for supplementing any formal investigation with non-formal considerations. In this sense the interrogative means are better suited to deliver these desiderata: allowing for a partial formal modeling and being suited to accommodate the informal parts. The models' context will provide the frame in which the nonformal considerations can be substantiated. Besides engaging interrogative tools for expressing modeling processes it also seems to be a more flexible way of forging their explanatory virtues.

Another accusation could draw on the fact that both central concepts – the explanation (conceived primarily as an answer to a why-question) and the model concept are too vague and in need of more disambiguation. But it is not the case for us that we are going to take one concept as *explicatum* for the explication of the other. Rather what I mean when invoking the coupling of modeling view and interrogative approach it is that this clarification should proceed in tandem and not in isolation. Instead of explicating the concepts in isolation we are seeking to identify and illuminate a process and its product that is partially captured by both concepts: explanation and model. What I'm pleading for is rather situating the explication of explanation in the frame of an investigation that takes modeling processes as essential for explanations.

Further on I would like to draw attention and discuss a well-articulated account under this view due to Hintikka. It seems to me that this account represents a successful example of a consistent application of interrogative ideas to the logical study of scientific explanatory enterprise. Most such intentions of application stop usually at the level of suggestion for an ambitious working agenda – as for example the case of Bromberger's ideas. The most promising characteristic of Hintikka's interrogative model of scientific inquiry is to be found in the way it recuperates the dynamics of the inquiry act, providing a natural frame to capture the processuality of the explanatory inquiry.

In his ’81 paper on the logic of interrogative model (IM shortly) of scientific inquiry, Hintikka announced already the relevance of his proposal for the problem of scientific explanation. In this context an explanation is not simply an answer to a why question; it is a process that provides the missing collateral information that makes the conditions of the conclusive answerhood of some questions satisfied. A more explicit articulation of the idea was provided in a recent paper, which will be discussed later in this chapter.

The IM conceives scientific inquiry as a game that the scientist plays against Nature. The inquirer subjects Nature to a series of questions; its answers together with the scientist’s background knowledge allow the inquirer to reach his goal by answering the initial question “C or non-C?” The process can be formally captured through a Beth tableau, which the inquirer seeks to close. He uses also besides deductive moves interrogative ones. A very important aspect of the model is the way it conceives the adequacy of the answer. A reply to be a real answer has to change the epistemic state of the inquirer. Hintikka makes recourse to epistemic logic, which is needed in order to spell out the semantics of the IM. We’ll see later, when discussing his latest contribution, how the epistemic aspect creates the interface for the junction with the modeling approach.

Regarding the history of the IM, the model did not surface as a main combatant in the explanation debate during the ‘80s. This is due probably to the reference to Hempel’s[88] model that wasn’t a too attractive option in the fourth decade of the debate. Partially it is also due to the fact that van Fraassen’s bold view based on interrogative approach obscured Hintikka’s variant. In the ’90s the IM started to be invoked more often in the context of the explanation topic, Sintonen being one arduous advocate of its use. Another articulated application of the IM is to be found in a paper in mid ’90s in which Weber provides us with an account on the process of constructing an explanation by directly applying the IM. His account centers on Hempel’s ideas, an explanation being conceived as a potential *explanans* plus a derivation of the law contained in the *explanans* from a general theory of laws. Due to its reliance on the Hempelian schema the account is limited and can be subjected to critiques analogous to the one against the DN model; but Weber’s account proves this way the generality and the flexibility of the IM.

*** * ***

In a recent paper (2005), Hintikka fully develops, together with Halonen, the old suggestion by explicitly articulating a model of scientific explanation, or rather a model for the process of explanation. The account is a well-articulated proposal, which brings together different ideas that were developed in Hintikka’s work over time. They are not only ideas around the IM of scientific inquiry, but also those developed under the topic of epistemic logic. I’ll further comment and evaluate the account in the light of the inquiry I’m pursuing. By drawing on the promising parts of the account, as well as exposing the

88 Hintikka considers rescue by viewing it as a specialized case of his IM schema.

limitations and commenting on the critiques that were formulated on the account, I'll try to show how the modeling view can make better sense of the elements which constitute the explanatory ingredients involved. The elements, which find a natural place in a modeling framework, will be especially emphasized. On the other side, possible ways in which IM can be useful for analyzing the explanatory virtues of models will also be suggested through this investigation. The modality of making IM specific to explanation through models should be revealed through the particularities of the modeling process. There is also a sense of explanation through models that seems to be most suited for an application of IM. Furthermore there is a sense caught in the idea of model building on which I'll discuss briefly later in the chapter.

To begin with: an explanation is according to Hintikka and Halonen a kind of derivation – not purely deductive – of the *explanadum* from the background theory. In order to do this, *ad explanandum* conditions have to be found; this input is provided through the answers 'Nature" gives us. The proposal is set under a new light by comparison with the other previous times when it was discussed ('81s, '99s) and by making explicit reference to the process of explanation Hintikka and Halonen emphasize the difference between background theory and covering law. By using Craig's interpolation lemma, they identify a formula that can play a role similar to the covering law in Hempel's model. The way the authors state this result is the covering law theorem. The deductive component in an explanation is identified to lie in the derivation of the covering law from the background theory. The three components: the covering law, the background theory and the *ad explanadum* conditions are to be found in any explanation, but they could play different roles. In some explanations the derivation of the covering law takes the first stage, while in others the finding of the *ad explanadum* conditions constitutes the most important step. In what regards the explication of the derivation, it is not possible to find general adequate rules to lead the inquirer from the background theory to the *explanadum*. But the IM of scientific inquiry provides the approximate language (through a sort of a bookkeeping method) that allows to capture formally and to keep track of the explanatory process.

Now from a modeling point of view, we can try to spell out here the structure of an explanatory model. Nothing seems to speak directly against a model that can harbor all the mentioned elements. Nevertheless, as I will point in the following paragraphs, the account is quite a limited one to be adequate for capturing explanatory models in general. Particular types of models could, in a more or less approximate way, embody the authors' ideas. Some of the ingredients appear to fit more naturally, as for example the *ad explanandum* conditions will match better the characteristics attributed to models, i.e., the specificity of the model to represent a particular situation. On the other side, the modeling process involves also a selection of theoretical elements that are integrated in the model in order to achieve the explanatory goal. The derivation step constitutes probably a too restrictive condition; it will be more adequate for a theoretical type of model. There is also a limitation to which Schurz points in his commentary that, besides the deductive and interrogative moves, abductive moves should be included also in an explanatory process.

The most important aspect that appears to be left out is the pragmatic dimension of the process. The account is developed in a syntactic and in a declared intended semantic register without allowing pragmatics to play any role. Pragmatics is explicitly mentioned in one place: related to the choice of the theoretical background. In what regards the usual place in which pragmatics is invoked in relation to the emphasis of the why-questions, the authors offer an alternative to the contrast class analysis claiming this way to avoid van Fraassen's radical conclusions. The problem is handled through the notion of queried constants. Schurz showed in his critique[89] of the account that the idea does not make good sense informally, and he provides us further with a solution that introduces extra presupposition – emphasis presuppositions. Nevertheless both ways to deal with this issue do not, in my opinion, cut out the role played by pragmatics at this point. Without invoking contrast classes, what determines which constants are queried or what emphasis presupposition the inquirer is willing to admit is in a good sense a matter of pragmatics.

The authors claim that what counts as the relevant background theory is a matter of pragmatics, meaning that the inquirer chooses whatever theoretical structures are at hand in order to fulfill the explanatory job. The background theory has to be taken for granted and not constructed as in the case of the covering law. It is also required that the background theoretical ideas are not *ad hoc* ones in the sense that they are tailored only for the specific purpose of the explanation sought. The most problematic point seems to be the fact that the background theory is not subjected to a building process.[90] It is not only the fact – which is very important – that an explanation involves modification or new construction of some theoretical items, but it seems to me that excluding this element runs contrary to the move to recuperate the process of explanation. This process also comprises the inquiries into the theoretical entities and how to make them adequate for the explanatory purpose pursued.

Hintikka and Halonen could be accused of developing a syntactic biased approach. In the response to their critiques the authors try to reveal also a semantic dimension of their account. They announce a new interpolation theorem – the explanatory interpolation theorem – which uses a simplified version of the Beth tableaux. The new interpolant formula I_L is the "real *explanans*" made explicit, instead of the old implicit interpolant H[b]. In the authors' notation having $\tilde{F} \square I_L$ and $\tilde{I_L} \square G$, where $\tilde{F} \square G$ is the conclusion, the I_L specifies the individual occurrence in the structures specified by F and the features of the structures that make models of F be models of G. A direct modeling interpretation of the Hintikka and Halonen's ideas could read I_L effectively as disclosing an explanatory model or some core of it – the authors' suggestions point in this direction too. I_L "literally lists different individuals […] plus the dependencies between them and other individuals that suffice to imply the *explanadum*". Hintikka is drawing on an earlier idea he developed by stressing the fact that the dependencies between different objects are reflected in the quantifiers of a formula and that the order of these quantifiers reflects directly the steps of the explanatory process. The covering law is seen in this sense as a summary of the entire process of explanation. We have at this point a simple sketch of a

89 Published in the same issue of Synthese.
90 Sintonen also points to this limitation.

possible explanatory model in a semantic perspective, in a sense. It makes some first steps on the way suggested by Sintonen, i.e., to use a semantic view (in particular the structuralist conception of theories), to provide the frame in which to further clarify and amplify the IM.

Other critical reactions to the account pointed to the issue of the relevance of the *explanans*. For Sintonen, for example, the account fails short of securing the relevance of the premises. It may be interpreted in the sense that the relevance quest gets split in two different ones: an explanatory relevance given through the background theory and one through the *ad explanandum* conditions, since both are to be chosen in order to establish the relevant explanatory relation. Some could claim that the difficulty of the problem is diluted by recasting it in more specific ways. To argue further along this line, that the explanatory relevance issue is dissolved or expedited to other less-problematic types of relevance is not a task that can be undertaken without problems. The most probable outcome is that it restates in other terms the initial relevance issue. Stretching a bit the context, we could see how the explanatory relevance issue could become meaningful in the frame of the unification approach in the case of the background theory and for the *ad explanandum* case in a causalist approach frame.

Nevertheless the problem of the relevance is claimed to be solved by the authors, and the solution is spelled out in the explication offered through the new interpolation theorem as mentioned in the previous paragraph. The interpolation statement "I_L tells what it is about the structures specified by F that makes the models of F to be also models of G". If our task is to "bring about G", knowing that G is logically implied by F does not suffice; knowing I_L is the key for it shows "what kind of structures we have to find or construct in order to realize G, in that it lists the individuals that have to exist and the general laws that bind them together which guarantee this result." This could be taken as a description of a model-oriented sort of explanation in a semantic register. One could further try to exploit the idea and couple it with such an approach as Hughes' on explanation through theoretical models, or that of Hartman-Frigg. In this sense one should watch for bridging the differences arising from taking the individuals and relations to be among various representations or between the idealized objects in the models.

An interesting distinction related to the IM that should be considered in the modeling context refers to two types of questions. The distinction is one between the big initial question for which the final answer is sought and the small operative questions whose answers make out the intermediary steps of the inquiry. A first move in our effort to apply the IM to model-based explanation will be to identify the big question with the major question the model is supposed to answer, but it does not seem obvious in general what such a why-question should be. Looking at some explanatory models usually considered as examples – take the model of ideal gas for instance – it seems that they can answer not only one specific why-question but many. They are a kind of reservoir in which we can generate answers for different why-questions. What specific questions they will answer, it will depend mostly on the context of inquiry. But this view assumes the model as an already built entity that can contribute towards extending further the scientific knowledge. How are we than to identify such a question for which the model, as an entire representation of a phenomenon, will count as an answer?

One way is to look for the question that gave the initial impulse for the modeling, the one that drove the process of the construction of the model. The model as the final product of the process should stay for the answer of this question. One problem is that such initial questions lack a clear defined identity. Such questions make indeed for the most interesting interrogations in science as Sintonen also emphasizes, triggering explanation searching processes but their "identity is initially indeterminate."[91] As we proceed in our inquiry the question get more precise. We encounter here an inadequacy of the IM model since it refers to why-questions that express requirements for evidence and are therefore precise defined, rather than explanation requirements.

An important observation that could open new perspectives in making use of interrogatives in the modeling process is the fact that the model as the final product seems to provide an answer to a how question. The explanatory model usually offers a description of the phenomena explaining to us how the things behave. A new strategy opens in which we could try to account for an explanatory process from an interrogative view by considering how an indefinite why-question gets transformed into a well defined how question and its corresponding answer. I'll not pursue here this line but I think it offers a promising way to approach the issue of explanation through models. It will be worth therefore to check its domain of validity.

Taking into account the previous discussion on applying the IM to model-based explanation and the open problems related to the identification of questions and the answers in the modeling context, a first pertinent step will probably be to limit it to episodes that constitute parts of the process instead of the entire process. The IM could account best for the partial moves in the modeling; the application to the entire process could involve a second level of IM, i.e., taking the final answers from the local episodes as the intermediary steps. But the main issue is to render the model here as an answer to the big initial question, as mentioned previously. The IM schema might be in need of transformation for this purpose.

Sintonen is worried further about another detail related to the IM model, namely, the fact that the conclusiveness conditions for the two different types of questions seem to be different. The conditions seem to be adequate for the small operative questions, since these questions ask for pieces of evidence. But for the big question, if it is to answer a why question, the condition proves to be inadequate. In order to see this distinction clearer I'll recall further shortly the epistemic aspects of the IM.

The originality of Hintikka's model among other interrogative approaches comes mainly from the epistemic considerations embedded into account. Making appeal to notions from epistemic logic, the epistemic aspect is captured by introducing the K-operator expressing what the inquirer knows and the /K notation to express the independence from a sentence-initial K. The why-questions are construed by taking into account the epistemic dimension in the following way. The desideratum of a question expresses the epistemic state of affairs, which the question is calculated to bring about. The reply is also conceived as being an epistemic state of affairs i.e., the one that is brought about by the inquirer's

[91] See Sintonen's critique to Hintikka and Halonen paper.

acceptance of a true response. But the reply does not automatically imply the *desideratum* (and so satisfy the inquirer request of information). A conclusiveness condition is stated to be satisfied in this sense. The knowledge expressed by this last condition has also a different status from the other knowledge. It is a different kind of information than the factual one contained in a simple reply. It is of conceptual nature rather than a factual one. According to the authors the conclusive conditions make sure that the inquirer knows what the reference of the term that makes out the answer is.

As Sintonen also rightly points, in the case of the why-questions the answerhood condition is fulfilled when the inquirer can say 'Now I know why X happened' in the sense that I understand why it happened. To invoke understanding as a necessary step in accounting for explanation increases the plausibility of the account. But the IM cannot accommodate understanding in its frame as Sintonen also noticed; a fact which discloses the limitations of the model. By coupling the IM with the modeling topic projected on a larger view of scientific inquiry, the understanding process could be better articulated as the answerhood condition.

The need for a broader view comes out also from the way the authors try to explicate the answerhood condition. In explicating the conceptual character of the knowledge involved in the condition they draw on the example of a controlled experiment. In such an experiment the dependence of a variable Y on another X is studied. The experimental data are plotted but such a graph does not suffice as a final answer. The authors point to the fact that the answerhood condition requires knowledge of the mathematical function that represents the graph. It is in this sense, as the authors emphasize, that the information is not empirical but conceptual.

The above example seems to me to show openly that the story behind is more complicated and IM is insufficient to account for it. The example would imply that an experimenter could not achieve understanding of the studied phenomenon without translating the results into a theoretical knowledge of some sort. This contradicts the natural intuition that understanding is locally sufficient and an experimenter could rest content with her results claiming understanding of the phenomenon for her level of inquiry without appeal to theories. It is for a theoretician's perspective that the data are in need of such a mathematical model, as a specific function capturing its graph. The identification of such a function would involve another additional constructive episode that should be also rendered through a similar IM model. Abductive moves would have a bigger role in this case. Besides it becomes clearer that the IM has to include also other sources of information than Nature only – as for example the other scientists follow.[92]

Viewed from a larger perspective on scientific inquiry the episode is in fact a case of embedding a scientific representation into a more theoretical one. The main idea of the structuralist conception of theories that conceives explanation as an embedding provides a proper way to capture this case. By considering understanding (i.e., providing an understanding) as an answerhood condition, the authors' characterization of the knowledge as conceptual can be also accounted for in a natural way. One way to get a better sense will be to view understanding as achieving coherence into a corpus of

[92] Sintonen also emphasizes this point.

knowledge (using for example Schurz' way of putting things). Taking explanation as the fitting of the *explanandum* by means of the *explanans* into the network of knowledge of the inquirer, the conceptual knowledge will reflect the theoretical engagements at work in this process. This could draw out further results from the unificationist register. Furthermore, framing the above issue into a modeling view calls on the characterization of understanding that I discussed in previous parts of my work. The local type of understanding that a model offers has to qualify further the understanding as an answerhood condition in the IM frame.

* * *

As it comes out from my previous presentation, one idea which seems to crop up constantly, if considering IM for the issue of explanation through models, is that the explanation process would be identifiable with a model building one (which I'll call the identification thesis). Some immediate precautions and specifications are in place here. The fact that my plea is not intended to be for a general theory of explanation through models was clear from the beginning. Qualifications through specific kinds of scientific modeling entities have to be central in the inquiry. But there is also a variation among the different ways we can claim explanation through models: for example as caught in the structure of the model or as caught by its dynamics. Accordingly we have to acknowledge the existence of different ways of pursuing the investigation. What is claimed in the identification thesis is contained mainly in the following idea: by identifyig the two processes we will have a consistent modality to capture a major source of the explanatoriness exhibited through models.

A parenthesis regarding the process approach: Sintonen points to a subtle distinction between the process of explanation and the process of seeking an explanation. He rightly emphasizes the fact that we are still in a rational reconstruction scenario when we take the first process as our object of analysis. In the discovery context the searching process should include all the dead ends and unanswered questions raised during our inquiry. The limits of our pretension to situate ourselves only in a genuine discovery context become clear. As mentioned in a previous section the classical distinction between the two contexts appears obsolete and a new view, which takes into account a larger variety of contexts, should be taken seriously into consideration. In my opinion a justified position will be to acknowledge the fact that the context is actually the one of an explanation in no need of any further reference to the classical distinction discovery–justification contexts. This way we can dismiss any serious consequence of Sintonen's distinction for our inquiry into the explanatory modeling.

One may further claim rightly that not all model-building processes are explanatory processes (only the ones that exhibit explicitly this purpose). But there are two ways of putting things under this issue: a pan-explanatory one and a restrictive one. The restrictive one will seek to establish the conditions meant to characterize the narrow sense of explanation through models concentrating on the models that are purely explanatory in nature (if such sort is to be found). From the methodological point of view it will deal more with theoretical models (since these are the best candidates for being primary explanatory models) in search for conditions. The possible drawbacks which can be identified from the beginning are related to the possibility of the derivation of the explanatory virtues from theories. Furthermore, the sense of explanation as present in other types of models than the theoretical ones runs the risk of being lost this way.

The pan-explanatory view takes the scientific inquiry more seriously as the basis for explanatory processes. The detour thesis finds a more important place here, since many other purposes of the scientific inquiry are means for the more fundamental explanatory one. In this sense the view allows the search for explanatory virtues for all types of models. The challenge is of course how to pick out the explanatoriness out of such diverse sorts of modeling? This explanatoriness quality has to be linked directly to the fundamental nature of the scientific inquiry. The pretension of a general account on scientific explanation can threaten the approach by bringing it back into the register of a classical approach on explanation. My suggestion is that we should keep this last strategy as a "regulatory idea", but retain the first one as our working option.

To conclude the discussion around the interrogative idea and its involvement into the approach on explanation through models, I'll recall the main points. The interrogative idea proved to be more flexible and generally applicable for the explanation topic than any other fundamental ideas as unification or causality. This generality is especially enhanced by the fact that scientific inquiry is essentially an interrogative process. Although interrogative means were used in the philosophy of science in the explanation subject, they are also adequate to address modeling issues. Modeling as a specific sort of scientific inquiry can be also subjected in a straightforward way to the application of the interrogative ideas for its explication.

As in the case of explanation, I think there are no real chances to develop a general account of modeling in a pure interrogative register. In the explanation issue, explanation could be naturally captured under an interrogative view as an answer to a why-question; even so the promises for a general model proved to be hopeless. In the model case, approaching it through an interrogative view could be a more demanding task since there is no obvious way to see how questions and answers can bear on directly and exhaust the issues posed in an inquiry of the modeling process. Addressing modeling issues seems to involve more elements that are not reducible entirely to interrogations. The way we can use the questioning-answering episodes is to give us only a partial clarification of the modeling process. It could contribute to the clarification of some elements of the modeling process. But the situation could be made more manageable if we take further into consideration the different types of models and the specific episodes of a modeling process. The questioning-answering episodes could be used to represent some of the steps during the modeling process – and in this sense the IM can find a direct application in some of such steps.

Nevertheless my main concern is not the articulation of an interrogative account of modeling but rather the inquiry of the explanatory virtues of models by using the interrogative ideas among others. In this sense I pointed out that both the interrogative type of approach on scientific inquiry and explanation and the modeling view on science exhibit similarities and move in the same direction. This fact alone should raise hopes for a successful implementation of interrogative means in the modeling topic.

The main advantage of making recourse to the interrogative ideas lies in the fact that it could provide the necessary intermediary terms and it enables us to bring closer the two topics of explanation and modeling. It provides the common ground for their most plausible articulation. Each of them can find more or less a separate articulation using interrogative means. This kind of articulation can be only a partial one, if we are to engage it into the inquiry of model-based explanation. I think that by addressing directly the topic of explanation through models under an interrogative view, we can hope to capture better the specificity of our problem.

Furthermore we can claim that the interrogative framework is necessary but not sufficient for our inquiry strategy. The necessity of appealing to this view is justified by the fundamental interrogative nature of scientific inquiry. But we should expect also the need for more constraints that are to be imposed from outside the interrogative frame in order to address our topic. The lack of sufficiency seems to be an obvious characteristic but skeptics could doubt the necessity required. In analogy with the critiques brought to the interrogative approach in the explanation debate they would invoke the weakness of the logic of interrogatives. But we do not intend a "rigorous" explication in terms a particular logic. Rather what I think to be required for our inquiry is a selective appeal to different conceptions and types of logic. Hintikka and Halonen's account integrates such elements from epistemic logic and interrogative logic and that was one reason why I discussed it in detail.

IM constitutes a well-articulated attempt of applying the interrogative elements to explicate the process of scientific inquiry in general. The 2005 account of Hintikka and Halonen seeks to extend the IM to the explanation problem. There are a number of limitations and distortions that the account suffers from as pointed out above. The IM proved to be in need of much modification and enrichment in order to be able to deal with the explanation topic. But what mainly counts here is that the possibility of addressing explanation under an interrogative_view centered on the process of scientific inquiry is a viable one. By making recourse to modeling elements we can further specify and cut out the vague points that obscure the explication of the process of explanation. In this sense, a model typology will specify, for example, more constraints on interrogatives. On the other side by engaging such interrogative means we secure a robust way of pinning down the explanatory virtues in a modeling process.

Final Remarks

Wrapping up and restating the main morals

To draw my inquiry to an end, I'll review the major findings of my discussion in the previous chapters and point to the main morals extracted and the perspectives open by them as a consequence.

In the first chapter I began my inquiry by proposing some distinctions, which were meant to induce a certain order among the diverse approaches to the explanation topic and at the same time to set the broad outlines for my further inquiry as a consequence of their evaluation. The distinction between global and local approaches and the one between static and dynamic extend already existing suggestions from the literature. The distinction between explanation as an application *versus* explanation as a construction is a kind of regulative one, intended to induce in an approximate way the direction of further inquiry by trying to avoid some undesirable attitudes towards the nature of the explanatory process. The difference between theory-driven and non-theory driven approaches on explanation brings into the explanation problematic a distinction used by the adherents of the model-oriented approach in the philosophy of science.

By discussing these distinctions, I introduced and sketched in a general way what I considered to be the most plausible type of approach in today's context of the philosophy of science on the topic of scientific explanation. This should be a dynamic, local, and non-theory driven type of approach. In the first chapter of my work I tried to identify the ingredients, which might constitute viable candidates for pursuing such an account. By putting my argument in a more general perspective I can say that it intends to move along the main ideas argued by Nick Huggett in his paper on local philosophies of science – it discusses localism with regard to the scientific explanation topic. The particular solution that I see for implementing such a local perspective couples the inquiry on scientific explanation with the one on scientific modeling.

In the second chapter I address in a straightforward manner the suggestion of realizing the type of approach I pleaded for in the first chapter, by setting the inquiry on scientific explanation in a framework that takes into consideration the subject of scientific models. As a first goal I proceeded to show that such an agenda is not an implausible one (as would have been regarded some time ago) and therefore I tried to dispel the potential inhibitions surrounding it.

Given the ways the topics were developed in the philosophy of science, I began first to show that the fact that they were kept apart is just an historical accident, the main cause being the bold influence of the received view conception on the working agenda of the philosophy of science in the 20th century. In a first historical detour I accounted for the origins and evolution of the philosophical reflection on scientific models, and I discussed how the "depreciative view" on models which originates in Duhem's conception of science and was continued by the logical-positivists, had an influential impact on the philosophical approach on models. But besides such a straight casting on the development of the topic, I emphasized the fact that the authors from the received view do not share always the same detailed position towards models and the way they can provide scientific explanations. I introduced a more detailed reflection on explanation through models by discussing Hempel's explicit position towards models and their explanatory virtues. I took his position as a reference for building a first reactive argumentation. I considered this to be a first necessary step in the restoration of the credibility of approaching explanation this way. In the light of the later advancements in the philosophy of science I argued that the hindrances and objections raised by Hempel are not justified. The models' "limitation in range", the "oversimplification" and the idealizations implicit in models, as well as in any other scientific theoretical construct do not constitute good grounds to reject the possibility of realization of explanation at models' level. In fact Hempel had a simplified view on these concepts and aspects of scientific knowledge, a conception that followed the lines of the received view.

I addressed also some questions concerning the "bearers" of explanation, as for example regarding the importance of having an account on the units of scientific knowledge that provide the explanations or the role they should play in the articulation of an approach on scientific explanation. The above discussion is made with reference to the classical candidates for this job: laws and theories, and with reference to the major accounts on explanations. My position backs the view that affirms the need of reference to scientific units of knowledge, but their articulation should be coupled with the explication of their explanatory virtues. I pointed to the fact that there is a tension between an explanation approach that draws on the classical, well-structured scientific entities (such as the laws and theories) and the one that tries to fulfill the need for more pragmatization. In my view the tension could be solved by taking models as the referential scientific entities.

By way of complementing the reactive argumentation, I further followed up several positive ways of implementing the local approach taking as points of departure already existing attempts. By discussing Morrison's account on models, I emphasized the fact that her approach, despite the drawbacks exhibited, represents a bold step forward in advancing the cause of explanatory models. I evaluated the possibility of an argumentation built around the idea of the autonomy of models that she draws on, emphasizing the fact that there are more registers in which this can be considered. Her way is based on the functionalist view she adopts, but it leaves open the issue of how the explanatory sort of autonomy builds from the autonomy in building and functioning of models. I discussed further the different modalities in which the explanatory autonomy could be claimed against other constructs that played the role of bearers of explanation in classical accounts.

In general Morrison expedites in the end a real solution for the model-based explanation by casting it as an issue related to representation through models. She reiterates this way a tendency manifested largely in the model-oriented camp. Her denial of the possibility of a general characterization of the nature of representation along with some suggested types of representations follows the directions of a local kind of approach I'm pleading for. The weak parts lie in the functionalist view she adopts as central for her inquiry. Her points on explanation through models remain mostly implicit and unsubstantiated. I'm pointing this way to the intrinsic limitations of pursuing the agenda on explanation through models by drawing only on pragmatic aspects.

Some positive incentives for pursuing the agenda are brought forth by showing how a consideration of the models might enhance our chances of successfully inquiring into the nature of scientific explanation. By drawing on the distinction material-formal conditioning required by any approach on explanation, I suggest the fact that models constitute the right type of scientific unit to capture properly both conditionings. The materiality of models as emphasized by some authors makes them distinctive from other scientific sort of constructs in this sense. This could be recast also in the way models recuperate the particularity of the phenomena by contextualizing the generality of theoretical constructs. Another attractive offer is to be found in the fact that the dynamical aspects of scientific inquiry may be naturally read out through the modeling processes. As argued, this type of dynamics is more adequate to reflect the explanatory virtues than other sorts of changes.

I ended the second chapter by reviewing the reticencies that were voiced and that could still worry a hardliner opponent. The danger of arbitrariness of modeling, i.e., of "free proliferation" in Duhem's words, is generated by the less constrained nature of the models. The plurality and variety of functions that models can fulfill might dilute the "explanatory effectiveness" of models. Another characteristic is to be found behind the "insufficiency" accusation that can be stated in the following way: at most one can grant the models the possibility of offering potential explanation but they lack the resources to transform them into an actual explanation. It is possible to articulate particular answers for the above worries but they have to be projected against a more general background. This general background restates the fact that these types of objections depend and reiterate the old requirements and standards from the received account on explanation; they could be overcome only by a fresh look on the explanation problem. In a broad sense this is mainly what the intention of my plea is throughout the chapters of this work.

For a consistent part of my investigation (Chapter 3) I focused on the concept of understanding and its relation to explanation in general, extracting the possible consequences for the local-dynamic type of approach. This move is driven by the attitude that challenges the explanation-centered kind of approach, which dominated the well-known major previous approaches; this attitude takes an explication of understanding as derivative from the one of explanation if not almost implicitly identical. In the traditional approaches a "taken for granted" identity between the two concepts was assumed. Focusing from the beginning on the classical approaches, I exposed the inadequacies of different existing concepts or rudiments of concepts of scientific understanding for an inquiry of understanding as given through models.

As I pointed out, Hempel's view of understanding as nomic expectability placed understanding either at the level of laws or made it inadequate as exhibited through the narrow class of predictive models. Kitcher and Friedman's accounts offer explicitly an account of understanding but their implementation in the modeling context has to play down the implicit global character they exhibit. Further on I tried to show that even apart from this aspect, there are inadequacies in using their approach to address understanding as delivered through models. Salmon's explicit reflection on different types of scientific understanding neglects any reference to models; nevertheless his causal type of explanation can be identified as a specific type of explanation through models that articulate causal mechanisms. Instead, Lambert-Schurz' account on scientific understanding exhibits a more flexible concept of understanding that raises hopes for an application in models; this is due to the fact that the account was not tailored on a specific type of scientific unit. Even so there are other characteristics that inhibit a straight application to models. I identify three main such limitations of their account: one due to the statement view adopted, one due to the requirement of "maximal information", and another one due to the neglect of pragmatics.

From the overview of the main accounts on scientific understanding, I concluded that there is a need to place the issue under a more generous frame, which should allow also a role for pragmatics, and recuperate important aspects of understanding – as the epistemic one. I further discussed several recent attempts which move in this direction. From their evaluation I marked the points that seem to constitute viable entries in my working agenda.

I discussed particularly in greater detail de Regt and Dieks' account on scientific understanding and the main critique against it due to Trout. The interest in this account is given by the fact that it is a pragmatic one and attempts to deal with the topic without any reference to scientific explanation. The account makes reference to theories as the main providers of scientific understanding, but the criteria could be transferred relatively easy to models' case. In fact models, as I argue, would make for a more adequate reference for the authors' proposal. Nevertheless a direct reading of the criteria will narrow the range to specific types of models, theoretical models among them.

Their main point of plurality of means through which understanding is realized corresponds directly to the need to capture the variety of modeling ways. The problematic part (not only for a modeling view) is related to the CIT criterion (or CIM if we will modify it), i.e., criteria of intelligibility of a theory/model. With reference to models and in the worse case, this will restrict us to a certain type that is best exemplified through the theoretical models in physics (e.g., the ideal gas model). For a more general characterization, the idea of understanding as ability to manipulate different scientific entities – that the authors also mention in their defense against Trout's critique – is more appealing, but it remains entirely uncontextualized.

Trout's critique of the understanding as conceived in the philosophical theories of scientific explanation has to be taken as a warning against the dangers that may be encountered when articulating an account on scientific understanding. The critique of understanding as overconfidence and hindsight makes good sense in a naturalist frame, but it is not clear how to identify them outside this frame even in the case of the classical accounts on scientific explanation. On these lines, the naturalists' contribution seems to be one of guarding against any possible deviations of the account; my suggestion is that such a check should be considered for any proposal on scientific understanding.

De Regt and Dieks' account restates, although in some other way, the point I made in Morrison's case regarding the limitation of a pure pragmatics-centered approach due to the lack of resources. De Regt's criteria provide a good approach for some special type of models. As I argue further in the chapter these criteria have to be embedded in a more generous view.

In an effort to enlarge the register of the approach on the relation explanation-understanding, I reviewed a series of distinctions that are taken to be fundamental in different approaches. So, by casting the relation explanation-understanding through a product-process distinction or qualifying it by the potential-actual distinction, we engage in specific ways of building the explication. We need a certain modality to evaluate and weigh the plausibility of adopting some strategies rather than others. I argued for the need of taking more closely into consideration the attempts of some epistemologists (such as Zabzeski, Kvavig, Riggs or Grimm) in order to draw on the understanding concept and its relation to knowledge. The hope is that the results from the epistemologists' camp would suggest what would be the minimum necessary basic assumptions needed to ground the investigation.

The points made on the explanatory model in the modelist accounts address important issues for the topic. The identification of different episodes in the modeling processes as articulated, for example, by M. Morgan raises the quest of different sorts of understanding. In this sense such a casting of understanding, as ability to manipulate, could make for a unifying concept and should be modulated after the type of objects that are manipulated. Most important we have to require an explicit criterion of manipulation. The above problems remain unaddressed by the modelists even by those who reflected explicit on the problem.

Conceiving understanding as ability to manipulate makes for a very general unificatory frame – unspecific for different types of understanding, overwriting even the distinction scientific *versus* non-science. A main question will concern the way we could induce some differentiations. One option will be according to the type of activity given through the kind of objects we manipulate. Unfortunately, it seems that this way fails to deliver the wanted result, since many sorts of objects are manipulated in different episodes of scientific inquiry during modeling processes. If we regard directly the manipulation of models, we bypass already the understanding through a model. We could try to invoke reference to understanding through theories or rather bigger units that involve manipulating models. If we regard instead the manipulation of concepts, we induce an indistinction between understanding through models and understanding through any other scientific entity. An escape from this situation could come from an appeal to purposes as follows: we gain scientific understanding when we achieve successfully the purpose for which the model was built. Nevertheless it is possible that due to the many different purposes we find ourselves in danger of allowing too many types of understanding. Nevertheless we still want to recuperate the understanding related to explanation; therefore we will have to invoke the specificity of understanding given through explanatory purposes.

As we already saw in the case of Morrison's approach, representation is the other concept invoked in relation to the possible different modalities of explanation and understanding. The main task in this context is to account for the way in which manipulatory activities are related to different types of representation and are to be distinguished. One strategy is to ground the differences in abilities in the specificity of representation. This track was not really pursued and elaborated despite being announced often by the adherents of model-oriented approaches. The reason for this situation, I think, lies in the fact that it is hard to see how to untangle the uses of different types of representations. In order to identify the tasks involved in following this direction, I point to the main open issues that have to be addressed. The core question would be whether building and manipulating different representations involve the manipulation of some common means or different ones**.** We see here how another level of entities – the constituents of representation are considered as the objects of manipulation. A notorious problem rises in this situation due to the fact that different types of representation could constitute a model.

In order to get on a more concrete track, I discuss in greater detail some specific accounts that advanced proposals on explanation through specific models, especially theoretical ones. Hughes' and Yi's accounts on explanation through theoretical models in general and in particular in condensed matter physics could be seen as possible exemplary_articulation that are to be extended to other types of models; meanwhile Hartmann-Frigg's account provides a possible general schema for explanation through models. I think that Hartmann-Frigg schema is the most promising attempt (despite the fact that is only a sketch) to describe in a more general way an approach to explanatory models. I think also that their point regarding the independence of such an approach from any theory of representation, though in need of more backing, is a right one.

The major morals which I extracted from the discussion of the modelists' approaches state first of all the possible existence of several sorts of understanding that have to be taken into consideration. Consequentially, the problem to be addressed concerns the relations between these different types and the quest of how they contribute to the understanding through models in general. One solution, as I pointed out, would be to identify one such moment as defining for the entire modeling process. I take Hartmann-Frigg's emphasis on the understanding to be gained at the moment of knowledge transfer as an exemplification of this solution. But I rather see a genuine solution to this issue coming through an appeal to the interrogative approach to explanation and in particular to some implementation of it as in Hintikka's IM model which I discussed in the following chapter. The distinction between the main big question and the small operative ones would render better a frame for an inquiry into the different moments of gaining understanding in a modeling process.

A further moral of the discussion rests with the fact that the ultimate justification of understanding involves an appeal to some ingredient from outside the explanation register, such as grasping or ability to manipulate or awareness of the explanation (Hughes). In my opinion this points clearly to the topics and results from the epistemology area, as suggested earlier.

In the fourth chapter I discussed in detail the different possible ways to cast the idea of locality, as they were suggested by different authors, but also as it is suggested by scientific practice. Locality as rendered through causal detailing, although it made for an appreciated solution for some time in the explanation debate, has its drawbacks, which were revealed by some authors (as Batterman and

Stevens more recently). My suggestion is that the requirement for locality finds a better casting by taking models as its natural expression. Models comprise the sort of locality that is relevant for investigating the scientific explanatory practices. This casting makes it possible to capture the old intuitions, as the causalist one, but also the ones linked to the pragmatic way of interpreting locality.

A contradiction that rises apparently when considering understanding at the model level is caused by the fact that understanding was placed usually in a global register. I considered first the coherence mark of understanding and the unification that was claimed to be realized through it. I analyzed the ways we could extend the already existing concepts of unification through analogy in the modeling context. Furthermore I pointed to the fact that although they seem to be inadequate to capture understanding through models, they could at most be taken as suggestions for an inquiry of this kind. By restating Morrison's thesis of the existence of various sorts of unification, some which can fail to exhibit explanatory qualities, I suggested that her conclusions seem to hold also in the context of explanatory modeling. It appears that such a general characterization as "unification in practice" which Morrison mentions as a specific variety of unification would refer directly to the sort of unification realized through models. In order to detail the idea behind this suggestion in a more consistent way I drew on Boumans' account on unification in mathematical models in economy.

I further argued for the main advantage and the perspective that Boumans' idea has for the approach on explanatory models. In this sense I pointed to the fact that viewing unification as integration of different ingredients fulfills the wish for a dynamic approach on unification whether it is qualified as explanatory or not. Of course in order to get the needed qualification we have to spell out the "explanatory requirements" which constitute only one item in the list of the adequacy criteria that Boumans states for the integration process. My suggestion is that the integration process should be qualified by the main purpose of the model building.

Of course, one can push further the demand for explicitation by asking for some clearly-cut formulated criteria in order to achieve a successful integration as reflected in the way the different ingredients are brought into a unity. Given the variety of the representations, this integration requires finding a solution to the quest of how to obtain a satisfactory match between different representations**.**

One important moral that I wanted to underline through my inquiry is the fact that by placing understanding in a local register, its main characteristics, such as coherence or unification, despite the fact that they were discussed in global terms, can make perfect sense and have to be recuperated in a local register. There are elements from the global characterization that could find a direct translation into the local one. I discussed such examples as Gisjbers' argumentation for local understanding and Boumans' account on model building as attempts to handle the above idea. An important point that is to be emphasized is that the localization of these traits requires specific qualification to be given through the purpose for which the model is built.

Another secondary moral that emerged from my discussion restates the thesis of the variety of types of understanding to be identified in the modeling context. It restates also the need to select between these types of understanding and from different models the explanatory ones.

As it is mentioned in many parts of my work, an essential reference in my enterprise is to take into consideration the purposes the models are built for. My main question refers to the consequences we could draw if we are to view explanation and understanding as specific purposes among others. A more concentrated reflection was undertaken to see how we should draw on this reference. I argued for the point that explanation can be conceived as being a fundamental purpose underlying any scientific inquiry process and therefore any modeling process, too. This comes primarily from the fact that the scientific activity is in an essential way explanatory-driven. As P. Lipton noticed even manipulatory or predictive scientific activities involve an explanatory detour. This could induce the tendency to adopt a pan-explanatory view, i.e., that all modeling is explanatory in a more or less general sense. This view contains its intrinsic dangers.

As a further second step I address the relation that explanatory purposes bear to other types of purposes, and I discuss particularly two such purposes which were viewed also as being sufficient general purposes to drive any scientific inquiry. The descriptive purpose could be claimed also to be a fundamental one and even to overwrite the explanatory one. Contrary to such a position as Ankeny's who identifies pure descriptive models as a pre-explanatory one, I argue for the fact that the models are to be conceived as explanatory-charged even in such cases as the one that Ankeny takes as reference in her exemplification. A first argument draws on the idea of placing the issue in a well-articulated frame of a conception of scientific knowledge such as the structuralist conception of science. A second argument looks at another philosophical position, which uses the same example – the C-elegans model – to articulate an approach on explanation in biological and medical sciences. The position allows us to claim explanatory virtues even for such pre-explanatory models as the wiring diagram identified by Ankeny as illustrative for her account. It is therefore only in a relative sense that we can identify such a pre-explanatory stage, depending on which episode we are regarding as explanatory. I further explore the ideas advanced by Schaffner with an eye on the special types of models that are developed in some interdisciplinary fields of bioscience. In discussing Schanffer's approach, my intention is to point also to the fact that a model oriented reading of his account makes for a more coherent perspective, avoiding the different drawbacks it exhibits in the present form. To realize this better, there is also a need for an enrichment of his approach with other elements especially from pragmatics. The main moral of the discussions emphasizes the fact that by focusing our attention on such particular models we can recuperate better the characteristics proper to models rather than to other scientific entities.

I finish the fourth chapter by discussing one of the main ideas that, I think, should be incorporated into the type of approach I'm pleading for. The interrogative idea is of essential importance for the approach, providing the necessary infrastructure for a more rigorous articulation of the processes that are constituent of an explanatory process. By drawing on this idea I make also indirectly a clear choice in favor of a classical tool used in the approaches on explanation.

I discuss first the status the interrogative approach had, showing how it granted high hopes at the end of the fourth decade of the debate. It was expected to fulfill the role of mediating and integrating the other central ideas on explanation. I think we should retain this promise, which actually rests mostly unfulfilled over the next two decades. I argue for the fact that its realization in a modeling context has better chances, due mainly to the localism of the model-oriented approach.

As a first step I identified the common points as well as the divergences between the interrogative approach and the model-oriented attitude in the philosophy of science, answering such questions as why such a combination was ignored and what are the points to which one can appeal from the other view in order to get more clarification. I addressed also the objection usually brought to the interrogative approach on explanation, i.e., that it provides a too weak a means for forging the explanatory issue, rejecting it by showing first that in general the critiques cannot be taken in a stronger sense and second that under a model-oriented setting in particular, they lose even more from their effectiveness.

I made concrete my discussion of the potential benefits of the interrogative view with reference to a well-articulated interrogative model on explanation, i.e., the one of Hintikka and Halonen. The more recent detailed articulation of the account (2005), although it exhibits limitations as reveled by various critiques, it discloses at the same time the possible problems that a local perspective, based exclusively on the interrogative idea, would encounter, and the possible ways to deal with them. The importance of the account for my agenda is revealed also by the fact that it intends to be an account on the process of explanation rather than as a product.

After presenting the new recasting of the IM model, the main lines of discussion are directed towards the possible applications and suggestions it may rise in a modeling context. Some of the elements allow a direct reading in a modeling context; nevertheless there are points that fundamentally hinder such an application. The dominant syntactic flavor of the approach, though augmented in the last version with some semantic items, is identified as a main hindrance. The approach barely mentions the pragmatic aspects of the scientific inquiry so central both to modeling and explanatory practices. This neglect is reflected in a more or less direct way in the critiques that the account received.

In a second stage the points revealed by these critiques (Sintonen, Schurz) are solvable by appealing to the means mentioned previously. In this sense the conclusive conditions for which Sintonen worries concerning the two distinct types of questions: the small operative ones and the big guiding question, could be properly addressed by drawing on the epistemological developments on the line of taking understanding as the answerhood condition.

One interrogative idea that seems to me to be of great value to be implemented in a modeling context is the distinction between the big question and the small operative ones. There are nevertheless issues raised by the identification criteria of such questions, which have to be addressed. Besides these, there are concerns (raised also by the critiques such as Sintonen) regarding the conclusiveness condition. Despite these worries, I think that the idea is worth to be worked out especially in the light of the fact that we can bring an epistemic view to bear upon some of the issues. A further, more consistent way of situating the issues is suggested by taking a semantic view of theories as the structuralist conception, as a general background of the problem.

Finally some auxiliary ideas are mentioned as potential viable ways of setting further the agenda of the approach, such as: the identification of the explanatory process with the model building or the distinction between the process of explanation and the process of seeking an explanation. Pursuing the lines of such distinctions is seen as dependant on the more general setting adopted, i.e., under a pan-explanatory view or a narrower one.

To conclude: the overall moral of my thesis is that given the actual landscape of the philosophy of science a promising way of attacking the explanation problem is to pursue it in a specific register of which characteristics I identified and tried to make explicit. As a more concrete way I see such an endeavor unfolding under a model-oriented perspective which would provide this kind of register. My conviction is that under this general configuration we can also hope that our desire for a more fruitful relation between scientific practice and philosophical reflection could be fulfilled.

.

References

Achinstein, P. (1968). *Concepts of Science. A Philosophical Analysis*. Baltimore: Johns Hopkins Press.

—— (1983). *The Nature of Explanation*. New York: Oxford University Press.

—— & O. Hannaway (Eds.), (1985). *Observation, Experiment and Hypothesis in Modern Physical Science*. Cambridge (Mass): M.I.T. Press.

Ankeny, R. A. (2000). “Fashioning Descriptive Models in Biology: Of Worms and Wiring Diagrams” in *Philosophy of Science*, 67(3 Supplement), S260-S272.

—— (2001). Model Organisms as Models: Understanding the 'lingua franca' of the Human Genome Project. *Philosophy of Science, 68*(3 Supplement), S251-S261.

Apostel, L. (1961). Towards the Formal Study of Models in the non-formal Sciences. In H. Freudenthal (Ed.), *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences* (pp. *1*-37). Dordrecht: Reidel.

Bailer-Jones, D. M. (1999). Tracing the Development of Models in the Philosophy of Science. In L. Magnani, N. Nersessian & P. Thagard, (Eds.), *Model-based Reasoning in Scientific Discovery* (pp. *23*-40). Dordrecht: Kluwer.

—— (2003). When Scientific Models represent. *International Studies in the Philosophy of Science, 17,* 59-74.

Balzer W., Moulines, C. U., Sneed, J. D. (1987). *An Architectonic for Science*: *The Structuralist Program*. Dordrecht: Reidel.

—— and Moulines, C. U. (1996) (Eds.). *Structuralist Theory of Science, Focal Issues, New Results*. Berlin: de Gruyter

Barnes, E. (1992). Explanatory Unification and the Problem of Asymmetry. *Philosophy of Science, 59,* 558-71.

Bartelborth, T. (1996). *Begrundungsstrategien: Ein Weg durch die analytische Erkenntnistheorie.* Berlin: Akademie Verlag.

—— (1996). Scientific explanation. In W. Balzer and C. U. Moulines (Eds.), *Structuralist Theory of Science, Focal Issues, New Results* (pp. 23-43). Berlin: de Gruyter.

—— (1999). Coherence and Explanations. *Erkenntnis, 50*(2-3), 209-224.

—— (2002). Explanatory Unification. *Synthese*, *130*(1), 91-107.

Batterman, R. (2000). A ‘Modern’ (Victorian) Attitude towards Scientific Understanding, *The Monist, 83*(2), 228 – 257.

—— (2002). *The Devil in the Details. Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford: Oxford University Press.

Belnap, N.D. & Steele, T.B. (1976). *The Logic of Questions and Answers*. New Haven: Yale University.

Bird, A. (2005). Explanation and Metaphysics. *Synthese, 143*(1-2), 89-107.

Bishop, M. A., & Trout, J. D. (2002). 50 Years of Successful Predictive Modeling should be enough:

Lessons for Philosophy of Science. *Philosophy of Science*, *69*(3 Supplement), 197-208.

Black, M. (1962). *Models and Metaphors. Studies in Language and Philosophy*. Ithaca, New York: Cornell University Press.

Blatt, M., Wiseman, S., Domany, E. (1996). Super-paramagnetic clustering of data. *Physical Review Letters 76*, 3251.

Bogen, J., Woodward J. (1988). Saving the Phenomena. *Philosophical Review, 97*, 303-352.

Boumans, M. (1999). Built-in Justification, In M. Morgan & M. Morrison (Eds.), *Models as Mediators. Perspectives on Natural and Social Science* (pp. 66-96). Cambridge: Cambridge University Press.

Braithwaite, R. (1953). *Scientific Explanation*. Cambridge: Cambridge University Press.

Brandom, R. (1985). Varieties of Understanding. In N. Resher (Ed.) *Reason and Rationality in Natural Science: a Group of Essays* (pp. 27-51). Lanham, MD : University Press of America.

Brittan, G. Jr., (Ed.). *Causality, Method and Modality: Essays in Honor of Jules Vuillemin The University of Western Ontario Series in Philosophy of Science, 48*. Dordrecht, Boston: Kluwer Academic Publishers.

Brody, B. A. (1970) *Readings in the Philosophy of Science*. Englewood Cliffs (N.J): Prentice Hall.

Bromberger, S. (1992). *On what we know we don't know: Explanation, Theory, Linguistics, and how Questions shape them.* Chicago and London: University of Chicago Press.

Bunzl, M. (1993). *The Context of Explanation*. Dordrecht: Kluwer Academic Publisher.

Callender, C., & Cohen, J. (2006). There is no Special Problem about Scientific Representation. *Theoria, 21:1*(55), 67-85.

Campbell, N. (1920). *Physics: The Elements*. Cambridge: Cambridge University Press. Reprinted as *Foundations of Science*. (1957). New York: Dover.

Carnap, R. (1939). Foundations of Logic and Mathematics. In O. Neurath, Ch. Morris and R. Carnap (Eds.), *International Encyclopaedia of Unified Science. Vol.*1 (pp. *139*-213). Chicago: University of Chicago Press.

—— (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Cartwright, N. (1983). *How the Laws of Physics lie*. Oxford: Oxford University Press.

—— (1989). *Nature's Capacities and their Measurement*. Oxford: Oxford University Press.

—— (1997). Models: The Blueprints for Laws. *Philosophy of Science*, *64*, 292-303.

—— Shomar, T., Suárez, M. (1995). The Toolbox of Science: Tools for the Building of Models with a Superconductivity Example. In *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 137-150.

—— (1999), *The Dappled World. A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

—— (1999). Models and the Limits of Theory: Quantum Hamiltonians and the BCS Model of Superconductivity. In M. Morgan & M. Morrison (Eds.), *Models as Mediators. Perspectives on Natural and Social Science* (pp. 241-281). Cambridge: Cambridge University Press.

Cat, J. (2007). The Unity of Science, *The Stanford Encyclopedia of Philosophy*, E. N. Zalta (Ed.), URL = http://plato.stanford.edu/archives/win2003/entries/models-science/

Cordero, A. (1991). Intelligibility and Quantum Theory. In C. Dilworth, (Ed.), *Idealization IV: Intelligibility in Science.* Amsterdam: Rodopi.

Cornwell, J. (2004). *Explanations: Styles of Explanation in Science.* New York: Oxford University Press.

Da Costa, N., French S. (2000). Models, Theories, and Structures: Thirty Years on. *Philosophy of Science* 67, Supplement, S116-127.

––– (2003). *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning.* Oxford: Oxford University Press.

DePaul M., Zagzebski L. T. (Eds.) (2003). *Intellectual Virtue: Perspectives from Ethics and Epistemology.* Oxford: Clarendon Press.

de Regt, H. W. (2004). Discussion Note: Making Sense of Understanding. *Philosophy of Science, 71*(1), 98-109.

––– (2006). Wesley Salmon's Complementarity Thesis: Causalism and Unificationism reconciled? *International Studies in the Philosophy of Science, 20*(2), 129-147.

––– & Dieks, D. (2005). A Contextual Approach to Scientific Understanding. *Synthese, 144*(1), 137-170.

––– & Dieks, D. (Ed.) (2005). *Philosophical Perspectives on Scientific Understanding,* booklet for the Conference on Perspectives on Scientific Understanding, Vrije Universiteit Amsterdam.

Dilworth, C. (Ed.) (1991). *Idealization IV: Intelligibility in Science.* Amsterdam: Rodopi.

Downes, S. (1992). The Importance of Models in Theorizing: A Deflationary Semantic View. In D. Hull et al. (Eds.), *Proceedings of the Philosophy of Science Association, Vol.1* (pp. 142-153). East Lansing: Philosophy of Science Association.

Duhem, P. (1954). *The Aim and Structure of Physical Theory*. Princeton: Princeton University.

Elgin, M. and Sober E. (2002). Cartwright on Explanation and Idealization. *Erkenntnis, 57*, 441-50.

Essler, W. K. (1975). Zur Topologie von Verstehen und Erklaren. *Grazer Philosophische Studien, 1*, 127-141.

Fetzer, J. H. (Ed.) (2000). *Science, Explanation, and Rationality: Aspects of the Philosophy of Carl G. Hempel.* New York: Oxford University Press.

Forge, J. (2002). Reflections on Structuralism and Scientific Explanation. *Synthese, 130*(1), 109-121.

Forster, Malcolm, and Elliot Sober (1994). How to Tell when Simple, More Unified, or Less Ad Hoc Theories will Provide More Accurate Predictions, *British Journal for the Philosophy of Science, 45,* 1-35.

French, S. (2003). A Model-Theoretic Account of Representation (or, I don't know much about art...but I know it involves Isomorphism). *Philosophy of Science, 70*(5), 1472-1483.

Freudenthal, H. (Ed.). (1961). *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences*. Dordrecht: Reidel.

Friedman, M. (1974). Explanation and Scientific Understanding. *Journal of Philosophy*, 71, 5-19.

Frigg, R. (2005). *Models, Explanation and Understanding.* Draft of the presentation at the 2005 Conference on Philosophical Perspectives on Scientific Understanding, Vrije Universiteit Amsterdam.

─── (2006), Scientific Representation and the Semantic View of Theories, *Theoria 55*, 37-53.

Gabbay, D. M., & Guenthner, F. (2001). *Handbook of philosophical logic* (2nd ed.). Dordrecht, Boston: Kluwer Academic Publishers.

Gabbay, D. M., Johnson, R. H., Ohlbach, H. J., & Woods, J. (2002). *Handbook of the Logic of Argument and Inference. The Turn towards the Practical.* New York: Elsevier.

Gärdenfors, P. (1980). A Pragmatic Approach to Explanation. *Philosophy of Science, 47*, 404–423.

Giere, R. (1988). *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press.

─── (1999). *Science without Laws*. Chicago: University of Chicago Press.

─── (2004). How Models are used to represent Reality. *Philosophy of Science 71, Supplement*. S742-S752.

Gijsbers, V. (2007). Why Unification is Neither Necessary Nor Sufficient for Explanation. *Philosophy of Science, 74*(4), 481-500.

Grimm, S. R. (2006). Is Understanding a Species of Knowledge?. *The British Journal for the Philosophy of Science, 57*(3), 515-535.

─── (2008). Explanatory Inquiry and the Need for Explanation. *The British Journal for the Philosophy of Science, 59*(3), 481-497.

Hacking, Ian (1983). *Representing and Intervening*. Cambridge: Cambridge University Press.

Halonen, I., & Hintikka, J. (1999). Unification–It's magnificent but is it Explanation? *Synthese, 120*(1), 27-47.

─── & Hintikka, J. (2005). Toward a Theory of the Process of Explanation. *Synthese, 143*(1), 5-61.

Harre, R. *An Introduction to the Logic of the Sciences.* New York: St.Martin's.

Hartmann, Stephan (1995), Models as a Tool for Theory Construction: Some Strategies of Preliminary Physics. In W. Herfel, W. Krajewski, I. Niiniluoto & R. Wojcicki (Eds.), (1995). *Theories and Models in Scientific Process. Poznan Studies in the Philosophy of Science and the Humanities 44* (pp. 49-67).Amsterdam: Rodopi.

─── (1996), The World as a Process. Simulations in the Natural and Social Sciences. In R. Hegselmann, U. Müller & K. Troitzsch (Eds.), *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View* (PP. 77-100). Dordrecht: Kluwer.

─── (1999), Models and Stories in Hadron Physics. In M. Morgan & M. Morrison (Eds.) (pp. 326-346). *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press..

─── (2001), Effective Field Theories, Reduction and Scientific Explanation', *Studies in History and Philosophy of Modern Physics* 32, 267-304.

─── (2006), Models in Science. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/win2003/entries/models-science/>

Hegselmann, R., Müller U. & Troitzsch K. (Eds.) (1996). *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*. Dordrecht: Kluwer.

Hempel, C. G. (1965). *Aspects of Scientific Explanation and other Essays in the Philosophy of Science.* New York: Free Press.

Herfel, W., Krajewski, W. Niiniluoto I. & Wojcicki R. (Eds.) (1995). *Theories and Models in Scientific*

Process. *Poznan Studies in the Philosophy of Science and the Humanities 44*. Amsterdam: Rodopi.

Hesse, Mary (1963). *Models and Analogies in Science*. London: Sheed and Ward.

–– (1974). *The Structure of Scientific Inference*. London: Macmillan.

Hintikka, J. (1981). On the Logic of an Interrogative Model of Scientific Inquiry. *Synthese, 47*(1), 69-83.

–– (1984). The Logic of Science as a Model-Oriented Logic. In P. D. Asquith & P. Kitcher (Eds.), *Proceedings of the 1984 Biennial Meeting of the Philosophy of Science Association.* East Lansing, Mich.: Philosophy of Science Association.

–– & Halonen, I. (1995). Semantics and Pragmatics for Why-Questions. *Journal of Philosophy, 92*(12), 636-657.

–– & Halonen, I. (1999). Interpolation as Explanation. *Philosophy of Science, 66*(3 Supplement), S414-S423.

–– , Halonen, I., & Mutanen, A. (2002). Interrogative Logic as a General Theory of Reasoning. In D. M. Gabbay, R. H. Johnson, H. J. Ohlbach, & J. Woods (Eds.), *Handbook of the Logic of Argument and Inference. The Turn towards the Practical.* New York: Elsevier.

–– & Halonen, I. (2005), Explanation: Retrospective Reflections. *Synthese, 143*(1-2), 207-222.

Hitchcock, C. (1995). Discussion: Salmon on Explanatory Relevance. *Philosophy of Science, 62*, 304-20.

Hughes R.I.G., (1993). Theoretical Explanation, *Midwest Studies in Philosophy, 18*(1), 132-153.

–– (1997). Models and Representation. *Philosophy of Science, 64, Supplement. Proceedings of the 1996 Biennial Meetings of the Philosophy of Science Association,* 325-336.

–– (1999). The Ising Model, Computer Simulation, and Universal Physics. In M. Morgan & M. Morrison (Eds.), *Models as Mediators. Perspectives on Natural and Social Science* (pp. 97–145). Cambridge: Cambridge University Press.

Humphreys, P. (1992). *The Chances of Explanation: Causal Explanation in the Social, Medical, and Physical Sciences.* Princeton: Princeton University Press.

–– (1993). Greater Unification equals Greater Understanding? *Analysis, 53*(3), 183-188.

–– (2000). Analytic versus Synthetic Understanding, in J. H. Fetzer (Ed.), *Science, Explanation, and Rationality: Aspects of the Philosophy of Carl G. Hempel.* New York: Oxford University Press.

–– (2002), Computational models. *Philosophy of Science, 69*(3 Supplement), S1-S11.

–– (2004), *Extending ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford: Oxford University Press.

Jones, T. (1995). Reductionism and the Unification Theory of Explanation, *Philosophy of Science, 62* (1), 21.

Jones, T. (1997). Unification, Reduction, and Non-Ideal Explanations, *Synthese, 112*(1), 75-96.

Keranen, J., & Salmon, W. (2005). Explanatoriness: Cause versus Craig. *Synthese, 143*(1-2), 125-147.

Kitcher, P. (1976). Explanation, Conjunction, and Unification, *Journal of Philosophy, 73*, 207-212.

–– (1981). Explanatory Unification, *Philosophy of Science, 48*, 507-531.

––– (1985). Two Approaches to Explanation. *Journal of Philosophy, 82*, 632-639.

––– (1989). Explanatory Unification and the Causal Structure of the World in P. Kitcher and W. Salmon (Eds.), *Scientific Explanation. Minnesota Studies in the Philosophy of Science, volume 13* (pp. 410-505). Minneapolis: University of Minnesota Press.

––– & Salmon, W. C. (Eds). (1989). *Scientific Explanation. Minnesota Studies in the Philosophy of Science, volume 13*. Minneapolis: University of Minnesota Press.

––– & Salmon, W. (1987). van Fraassen on explanation. *Journal of Philosophy, 84*, 315-330.

Koertge, N. (1992). Explanation and its Problems, *British Journal for the Philosophy of Science, 43*(1), 85-98.

Knowles, D. (Ed.). (1990). *Explanation and its Limits,*. Cambridge. New York: Cambridge University Press.

Knuuttila, T. (2005). Models, Representation, and Mediation, in *Philosophy of Science, 72*(5), 1260-1271.

––– & Voutilainen, A. (2003). A Parser as an Epistemic Artifact: A Material View on Models. *Philosophy of Science, 70*(5), 1484-1495.

Kuipers, T. A. F., & Wiśniewski, A. (1994). An Erotetic Approach to Explanation by Specification, *Erkenntnis, 40*(3), 377-402.

Lambert, K. (1990). On whether an Answer to a Why-Question is an Explanation if and only if it yields Scientific Understanding, in G. Brittan, Jr., (Ed.), *Causality, Method and Modality: Essays in Honor of Jules Vuillemin The University of Western Ontario Series in Philosophy of Science, 48* (pp. 125-142). Dordrecht, Boston: Kluwer Academic Publishers.

Laymon, R. (1985), Idealizations and the Testing of Theories by Experimentation, In P. Achinstein and O. Hannaway (Eds.), *Observation, experiment and hypothesis in modern physical* science (pp. *147*-173), Cambridge, Mass. M.I.T. Press.

Leatherdale, W. (1974). *The Role of Analogy, Model, and Metaphor in Science,* Amsterdam, Oxford: North-Holland.

Leplin, J. (1980). The Role of Models in Theory Construction, in: T. Nickles (Ed.) *Scientific Discovery, Logic, and Rationality* (pp. 267-284). Dordrecht: Reidel.

Lipton, P. (2004). *Inference to the Best Explanation (second edition)*. New York: Routledge.

––– (2004). What Good is an Explanation? In J. Cornwell *Explanations: Styles of explanation in* science (pp. 1-22). New York: Oxford University Press.

Magnani, L. (2001). *Abduction, Reason, and Science: Processes of Discovery and Explanation*. Dordrecht: Kluwer Academic Publishers

––– & Nersessian, N. (Eds.). (2002), *Model-based Reasoning: Science, Technology, Values*. Dordrecht: Kluwer.

––– & Thagard, P. (Eds.). (1999). *Model-based Reasoning in Scientific Discovery.* Dordrecht: Kluwer.

McMullin, E. (1968). What Do Physical Models Tell Us? In B. van Rootselaar and J. F. Staal (Eds.). *Logic, Methodology and Science III.* (pp. 385-396). Amsterdam: North Holland.

––– (1985). Galilean Idealization. *Studies in the History and Philosophy of Science 16*, 247-73.

Morgan, M. (1999). Learning from Models, In M. Morgan & M. Morrison *Models as Mediators.*

Perspectives on Natural and Social Science (pp. 347-388). Cambridge: Cambridge University Press.

––– and M. Morrison (1999). Models as Mediating Instruments. In M. Morgan & M. Morrison *Models as Mediators. Perspectives on Natural and Social Science* (pp. 10-37). Cambridge: Cambridge University Press.

––– and Morrison M. (Eds.) (1999). *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.

Morrison, M. (1998). Modeling Nature: between Physics and the Physical World. *Philosophia Naturalis 35,* 65-85.

––– (1999) Models as Autonomous Agents, in M. Morgan & M. Morrison *Models as Mediators. Perspectives on Natural and Social Science* (pp. 38-65). Cambridge: Cambridge University Press.

––– (2000), *Unifying Scientific Theories*. Cambridge: Cambridge University Press.

Moulines, C. U. (2002). Introduction: Structuralism as a Program for Modelling Theoretical Science. *Synthese, 30*(1), 1-11.

––– & Polanski, M. (1996). Bridges, Constraints, and Links, in W. Balzer & C.U. Moulines (Eds.), *Structuralist Theory of Science: Focal Issues, New Results* (pp. 219-232). New York: de Gruyter.

––– (2005). Models of Data, Theoretical Models, and Ontology. A Structuralist Perspective, in Michael H.G. Hoffmann, J. Lenhard & F. Seeger *Activity and Sign* (pp. 325-333). Berlin: Springer.

Müller, T., & Müller, H. (2003). *Modelling in Natural Sciences: Design, Validation, and Case Studies.* Berlin: Springer.

Muschik, W. (1998). Experiments, Models, and Theories. Comment on Margaret Morrison. *Philosophia Naturalis, 35*(1), 87-93.

Nagel, E. (1979). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Indianapolis, Cambridge: Hackett Publishing.

Neurath, O., Morris Ch., & Carnap R. (Eds.) (1939). *International Encyclopaedia of Unified Science. Vol.*1. Chicago: University of Chicago Press.

Newton-Smith, W.H. (2001) Explanation, In W.H. Newton-Smith, (Ed.) *A Companion to Philosophy of Science*, Oxford: Blackwell Publishers.

Niiniluoto, I., Sintonen, M., & Wolenski J. (Eds.). (2004). *Handbook of Epistemology* Dordrecht, Boston: Kluwer Academic Publishers.

Pitt, J. (Ed.) (1988). *Theories of Explanation*. New York: Oxford University Press.

Psillos, S. (2002). *Causation and Explanation.* Montreal & Kingston, Ithaca: McGill-Queen's University Press.

Railton, P. (1978). A Deductive-Nomological Model of Probabilistic Explanation. *Philosophy of Science, 45*(2), 206.

––– (1981). Probability, Explanation, and Information. *Synthese, 48*(2), 233-256.

Redhead, M. (1980). Models in Physics, *The British Journal for the Philosophy of Science, 31*(2), 145-163.

Rescher, N. (1970). *Scientific Explanation.* New York: Free Press.

—— (Ed.) (1985). *Reason and Rationality in Natural Science: a Group of Essays* (pp. 27-51). Lanham, MD: University Press of America.

—— (2000). *Inquiry Dynamics,* New Brunswick, London: Transaction Publishers.

Rheinberger, H. J. (1992). *Experiment-Differenz-Schrift. zur Geschichte epistemischer Dinge,* Marburg: Basilisken-Presse.

Riggs, W. (2003). Understanding Virtue and the Virtue of Understanding. In M. DePaul, L. T. Zagzebski (Eds.) *Intellectual Virtue: Perspectives from Ethics and Epistemology* (pp. 203-226). Oxford: Clarendon Press.

Ruben, D. H. (1993). *Explanation,* Oxford, New York: Oxford University Press.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World,* Princeton: Princeton University Press.

—— (1998). *Causality and Explanation,* New York: Oxford University Press.

—— (1989). *Four decades of Scientific Explanation,* Minneapolis: University of Minnesota Press.

—— & Kitcher, P. (Eds.) (1989). *Minnesota Studies in the Philosophy of Science, vol. 13: Scientific Explanation*. Minneapolis: University of Minnesota Press.

Schaffner, K. F. (1993). *Discovery and Explanation in Biology and Medicine,* Chicago: University of Chicago Press.

—— (1998). Model Organisms and Behavioral Genetics: A Rejoinder. *Philosophy of Science, 65*(2), 276-288.

Schurz, G. (1995). Scientific Explanation: A Critical Survey. *Foundations of Science, 1*(3), 429-465.

—— (1999). Explanation as Unification, *Synthese, 120*(1), 95-114.

—— (2005). Explanations in Science and the Logic of Why-Questions: Discussion of the Halonen-Hintikka Approach and alternative Proposal, *Synthese, 143*(1-2), 149-178.

—— (Ed.) (1990). *Erklaren und Verstehen in der Wissenschaft,* Oldenbourg: Scientia Nova.

—— & Lambert, K. (1994). Outline of a Theory of Scientific Understanding. *Synthese, 101*(1), 65-120.

Scriven, M., 1962, Explanations, Predictions, and Laws. In H. Feigl, G. Maxwell (Eds.) *Minnesota Studies in the Philosophy of Science vol 3, Scientific explanation, space, and time* (pp.170-230). Minneapolis: University of Minnesota Press.

Simpson, J (2006). Simulations are not Models. In *Conferences and Volumes: [2006] Models and Simulations (Paris, 2006).* Retrieved from PhilSci Archive database.

Sintonen, M. (1984). On the Logic of Why-Questions, In P. D. Asquith & P. Kitcher (Eds.), *Proceedings of the 1984 Biennial Meeting of the Philosophy of Science Association* (pp. 168–176). East Lansing: Philosophy of Science Association.

—— (1984). The Pragmatics of Explanation, *Acta Philosophica Fennica, 37*

—— (1989). Explanation: In Search of the Rationale, in P. Kitcher and W. Salmon (Eds.), *Scientific Explanation. Minnesota Studies in the Philosophy of Science, volume 13,* (pp.253–282). Minneapolis: University of Minnesota Press.

—— (1990). How to put Questions to Nature, in D. Knowles (Ed.), *Explanation and its Limits* (pp. 267-

284), Cambridge, New York: Cambridge University Press.

—— (1996). Structuralism and the Interrogative Model of Inquiry, in W. Balzer & C.U. Moulines (Eds.), *Structuralist Theory of Science: Focal Issues, New Results.* New York: De Gruyter.

—— (Ed.) (1997). *Knowledge and Inquiry: Essays on Jaakko Hintikka's Epistemology and Philosophy of Science.* Amsterdam: Rodopi.

—— (1997). Explanation: The Fifth Decade. In M. Sintonen, *Knowledge and Inquiry: Essays on Jaakko Hintikka's Epistemology and Philosophy of Science* (pp. 225–238). Amsterdam: Rodopi.

—— (1999). Why Questions, and why just Why-Questions? *Synthese, 120*(1), 125-135.

—— (2005). Scientific Explanation: Conclusiveness Conditions on Explanation-Seeking Questions. *Synthese, 143*(1-2), 179-205.

—— and Kiikeri, M. (2004). Scientific Discovery, In I. Niiniluoto, M. Sintonen, J. Woleński. (Eds.), *Handbook of Epistemology* (pp. 205–253). Dordrecht: Kluwer Academic Publishers.

Sneed, J. D. (1979). *The Logical Structure of Mathematical Physics,* Dordrecht: Kluwer Academic Publishers.

—— (1994). Structural Explanation, in P. Humphreys (Ed.), *Patrick Suppes: Scientific Philosopher* (pp. 195-218). Dordrecht: Kluwer Academic Publishers.

Sober, E. (1983). Equilibrium Explanation. *Philosophical Studies, 4,* 201-210.

Spohn, W. (1991). A Reason for Explanation: Explanations provide Stable Reasons, in W. Spohn, B. C. van Fraassen & B. Skyrms (Eds.), *Existence and Explanation: Essays presented in honor of Karel Lambert.* Dordrecht, Boston: Kluwer Academic Publishers.

—— , van Fraassen, B. C., & Skyrms, B. (Eds.). (1991). *Existence and Explanation: Essays presented in honor of Karel Lambert.* Dordrecht. Boston: Kluwer Academic Publishers.

Stegmüller, W. (1975). Structures and Dynamics of Theories. *Erkenntnis, 9*(1), 75-100.

—— (1979). *The Structuralist View of Theories: A Possible Analogue of the Bourbaki Programme in Physical Science.* Berlin, New York: Springer-Verlag.

—— (1983). *Probleme und Resultate der Wissenschaftstheorie und analytischen philosophie.* Berlin, New York: Springer-Verlag.

Sterrett, S. G. (2002). Physical Models and Fundamental Laws: Using one Piece of the World to tell about Another. *Mind & Society, 3*(1), 51-66.

—— (2006). Models of Machines and Models of Phenomena. *International Studies in the Philosophy of Science, 20*(1), 69-80.

Steup, M. (Ed.). (2001). *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue.* New York: Oxford University Press.

Strevens, M. (2004). The Causal and Unification Approaches to Explanation Unified-Causally. *Nous, 38*(1), 154-176.

Suárez, M. (1999). The Role of Models in the Application of Scientific Theories: Epistemological Implications, in M. Morgan & M. Morrison *Models as Mediators. Perspectives on Natural and Social Science* (pp. 168–195). Cambridge: Cambridge University Press.

—— (2003). Scientific Representation: Against Similarity and Isomorphism. *International Studies in the Philosophy of Science, 17*(3), 225-244.

––– (2005). The Semantic View, Empirical Edequacy, and Epplication. *Critica: Revista Hispanoamericana De Filosofia, 37*(109), 29-63.

Sun R. (Ed.). (2008). *The Cambridge Handbook of Computational Psychology*. Cambridge: Cambridge University Press.

Suppe, F. (Ed.) (1977). *The Structure of Scientific Theories.* Urbana and Chicago: University of Illinois Press.

––– (1989). *The Semantic Conception of Theories and Scientific Realism,* Urbana and Chicago: University of Illinois Press.

Suppes, P. (1962). Models of Data. in E Nagel, P Suppes, A Tarski *Methodology and Philosophy of Science: Proceedings of the 1960 International Congress* (pp. 252–261). Stanford: Stanford University Press.

––– (1999). *Introduction to Logic.* Mineola, New York: Dover Publications.

Thagard, P. (1992). *Conceptual Revolutions.* Princeton: University Press.

––– & Litt, A. (2008). Models of Scientific Explanation, in R. Sun (Ed.), *The Cambridge*

Handbook of Computational Psychology (549-564). Cambridge: Cambridge University Press.

Teller, Paul (2001). Twilight of the Perfect Model. *Erkenntnis, 55*, 393-415.

Toulmin, S. E. (1961). *Foresight and Understanding: An Enquiry into the Aims of Science.* New York: Harper Torchbooks.

Trout, J. (2002). Scientific Explanation and the Sense of Understanding. *Philosophy of Science, 69*(2), 212-233.

––– (2005). Paying the Price for a Theory of Explanation: de Regt's Discussion of Trout. *Philosophy of Science, 72*(1), 198-208.

Tuomela, R. (1980). Explaining Explaining. *Erkenntnis, 15*(2), 211-243.

van Fraassen, B. (1985). Salmon on Explanation. *Journal of Philosophy, 11*, 639-651.

––– (1980). *The Scientific Image.* New York: Oxford University Press.

––– (1990). *Laws and Symmetry.* New York: Oxford University Press.

Vossenkuhl, W. (1995). Analyse und Hermeneutik. über "Verstehen". *Philosophia Naturalis, 32*(2), 295-317.

Wartofsky, M. W. (1979). *Models: Representation and the Scientific Understanding. Boston Studies in the Philosophy of Science vol. XLVIII.* Dordrecht: Reidel.

Weber, E. (1996). Explaining, Understanding and Scientific Theories. *Erkenntnis, 44*(1), 1-23.

––– (1999). Unification: What is it, how do we reach and why do we want it? *Synthese, 118*(3), 479-499.

––– & Van Dyck, M. (2002). Unification and Explanation: A Comment on Halonen and Hintikka, and Schurz. *Synthese, 131*(1), 145-154.

Weinert, F. (1999). Theories, Models and Constraints. *Studies in History and Philosophy of Science, 30*(2), 303-333.

Weisberg, M. (2007). Who is a Modeler? *The British Journal for the Philosophy of Science*, *58*(2), 207-

233.

Wiśniewski, A. (1995). *The Posing of Questions: Logical Foundations of Erotetic Inferences.* Dordrecht: Kluwer Academic Publishers.

Woodward, J. (1989). Data and Phenomena, *Synthese*, *79*(3), 393-472.

—— (1989), The Causal/Mechanical Model of Explanation, in P. Kitcher and W. Salmon (Eds.), *Scientific Explanation*. *Minnesota Studies in the Philosophy of Science vol. 13,* (357-383). Minneapolis: University of Minnesota Press.

—— (2000), Explanation and Invariance in the Special Sciences, *British Journal for the Philosophy of Science, 51*, 197-254.

—— (2000). Data, Phenomena, and Reliability. *Philosophy of Science. 67(S1),* 163-179.

—— (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press

—— (2006). Scientific Explanation. In E. N. Zalta (Ed.). *The Stanford Encyclopedia of Philosophy*. URL = <http://plato.stanford.edu/archives/win2003/entries/models-science/>

Yi, S. W. (2002). The Nature of Model-Based Understanding in Condensed Matter Physics. *Mind and Society: A Journal of Cognitive Studies in Economics and Social Sciences, 3*(5), 81-91.

Zagzebski, L. (2001). Recovering Understanding, in M. Steup (Ed.), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue* (pp. 235-252). New York: Oxford University Press.

Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. URL = <http://plato.stanford.edu/archives/win2003/entries/models-science/>