EXPLANATION THROUGH
SCIENTIFIC MODELS:
REFRAMING THE EXPLANATION TOPIC*

Richard DAVID-RUS

ABSTRACT: Once a central topic of philosophy of science, scientific explanation attracted less attention in the last two decades. My aim in this paper is to argue for a new sort of approach towards scientific explanation. In a first step I propose a classification of different approaches through a set of dichotomic characteristics. Taken into account the tendencies in actual philosophy of science I see a local, dynamic and non-theory driven approach as a plausible one. Considering models as bearers of explanations will provide a proper frame for such an approach. In the second part I make some suggestions for a working agenda that will further articulate a sketchy account of explanation through models proposed by Hartmann and Frigg.

KEYWORDS: scientific explanation, scientific models, understanding

I. The explanation topic and its fate

It is hard to overlook the status of the topic of scientific explanation in philosophy of science. The topic was one of the central subjects for few decades during the second half of the last century and it concentrated the efforts of some of the best philosophers. Salmon’s book *Four Decades of Scientific Explanation* documents the dense debate from its inception to the end of the fourth decade, i.e., the end of the eighties. Salmon – himself a major contributor to the debate – expresses at the end of the book his view for the future fate of the topic. He advances a complementary view in which the two major approaches that dominated the

* ACKNOWLEDGEMENT: This paper was made within *The Knowledge Based Society Project* supported by the Sectorial Operational Program Human Resources Development (SOP HRD), Financed by the European Social Fund and by the Romanian Government under the contract POSDRU/89/1.5/S/56815.

agenda at that time – the unificationist (as appears in Kitcher’s account\(^2\)) and the causalist conceptions of explanation (as expressed mainly in Salmon’s work\(^3\)) should complement each other. The pragmatic approach as advanced in van Fraassen’s\(^4\) account would play in Salmon’s view the decisive role of choosing in a specific context between the two previous complementary conceptions.

A quite immediate question that rises looking retrospectively would ask how much of this view got fulfilled in the meantime? Is it the case that the subsequent approaches focused their main effort in implementing such a kind of strategy? The obvious answer is no.

On the other hand there is a sort of fatigue that gradually made its way in the research attitudes in philosophy of science finding its expression through a sort of rejection of the subject of scientific explanation. Though officially a major topic, it does no longer concentrate the attention of philosophers as in previous decades. Nonetheless at the same time, one can notice a tendency that seeks to resettle the issue in new contexts of inquiry. As the philosophical research reveals new areas of interest in the philosophy of science, the explanation topic gains new dimensions challenging the old requirements and assumptions of the topic.

I’ll start my paper by first proposing a sort of ordering of the main approaches on explanation. This way I’ll argue for a type of approach that I see to be promising in the context of the recent tendencies in philosophy of science. In a subsequent step I’ll argue in favor of adopting a modeling view as a plausible framework to resettle some of the questions related to the explanation topic. I’ll continue by presenting briefly some proposals that go some way in the same direction as the one proposed. A specific account by Roman Frigg and Stephan Hartmann will be seen as the most viable frame for the inquiry into explanatory models. I’ll suggest some further steps that are to be taken in order to gain a more concrete articulation of this approach.

II. Getting some clues for a plausible approach to explanation

In this section I’m going to suggest four pairs of distinctions that will help me to characterize broadly the existing approaches on scientific explanation and will


\(^3\) As presented in his papers collected in the volume Wesley Salmon, *Causality and Explanation* (New York: Oxford University Press, 1998).

permit me further to identify a plausible direction that we could follow in our inquiry. I’m suggesting the following distinctions between: global versus local, dynamic versus static, theory-driven versus non-theory-driven and the approaches that have the tendency to view explanation as an application versus those approaching it as a construction.

The first distinction 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 therefore are to be used in a conception of explanation. A kind of global-type approach will be one in which the global considerations are viewed as central. This does not mean that only explanations that 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. Examples of such global approaches are Friedman’s⁵ or Kitcher’s accounts in which the quality of explanatoriness emerges at a global level – those of corpuses of knowledge (as in Kitcher’s account) or is given through reduction relations among laws as for Friedman. The accounts of Schurz⁶ or Bartelborth⁷ are more recent examples of this type of approach. Examples of local types of approaches are such as Salmon’s account in which explanation is understood as disclosing the local causal network that brings about the explanandum or the one of van Fraassen in which contextual factors are ultimately determinant for any explanation.

The second distinction is one between static and dynamic approaches. The distinction is primarily between approaches that take into consideration temporal aspects of the scientific structures versus those that ignore them. Or in other words, a dynamic perspective would consider the ways elements of scientific knowledge are modified or new elements are constructed in the process of providing an explanation for a phenomenon. The existing accounts are in general of the first sort. Kitcher’s account is one that addresses and integrates the temporal aspects of scientific knowledge. His conception considers different corpuses of knowledge from different historical periods as determining the explanations accepted as valid at that time. The account captures the dynamics at the macro level of scientific activity. Nevertheless, this sort of dynamics is only a specific one

---

Richard David-Rus

and not the relevant one for an adequate local solution on explanation. Kitcher provides us with 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 would be to provide some insight into how different elements of the patterns evolve, i.e., are chosen, modified or dropped in the course of searching for an explanation of a phenomenon while the macro-constraints at the level of the corpuses would partially influence these processes. In this sense a proper dynamic account would unfold at a local level.

The third sort of distinction between theory-driven and non-theory-driven approaches reiterates the lines of a reaction\(^8\) in today’s philosophy of science. In the ‘classical’ philosophy of science theories played a central role not only as a focus of investigation into their structure but also as making any other topic more or less dependent of such a solution. In the last decades this theory-centered view was shaken due to various investigations into what were considered to be aspects of secondary importance. Such secondary topics were the experimental activity or the modeling one and their related 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”\(^9\) influence. Nevertheless we could read out influences of the theory-centered but also signs from the opposite attitude in today’s approaches on explanation.

The last distinction advanced – the one between explanation as an application versus explanation as a construction – is rather a regulative one, that should direct our search under a certain perspective. The distinction points to the relation between the events (or facts) to be explained and the units used to carry out the explanation in the way it assumes in the background the view of the

\[8\] The reaction of the recent modelists, and I borrow here the expression “modelists” from Carlos U. Moulines, who used it in his book La Philosophie des Sciences. L’Invention d’une Discipline (Paris: Edition rue d’Ulm, 2006) to name the approaches in philosophy of sciences that developed model-centered conceptions of scientific knowledge

\[9\] I use here Suppe’s expression from Frederick Suppe, “The Search for Philosophical Understanding of Scientific Theories,” in The Structure of Scientific Theories, ed. Frederick Suppe (Urbana: University of Illinois Press, 1977), 3-232, to refer mainly to the logical empiricist conception of science.
explanandum entities that are to be plugged into formal schemas or patterns given by the theoretical constructs. The last view might be described using Cartwright words as “the vending machine” view. Her critique\(^{10}\) points to the fact that theories are conceived as one will “feed them an input in certain prescribed forms for the desired output” and after a while “it drops out the sought-for representation […] fully formed.”\(^{11}\) On the other hand, by using the notion of ‘construction’ I want to emphasize the different scientific activities involved in the process of explanation. Explanation as ‘construction’ takes seriously the idea that the representation of phenomena must be constructed and it is through such a process that we may get an explanation.

Now having laid out a sort of a grid on which we might compare different approaches I think that one sort of approach that we could consider not only as pertinent but also as promising for a future working agenda is a local, dynamic, non-theory driven kind and which will avoid conceiving explanation as an application. My choice is backed up by the conviction that such an approach will provide us with new valuable insights into how science works and will bring us closer to a more adequate consideration of scientific practice. To be more concrete I see some major benefits resulting from such an approach and that would justify the above choice. First of all, I think that such an approach would bring us closer to scientific practice through the fact that it emphasizes the local character and the dynamic aspects of explanation. Secondly, placing the issue in a dynamic frame makes this way room for the possibility of taking into consideration seriously the process of explanation (or conceiving explanation as a process rather than as a product). Besides this, from a larger perspective, it opens the possibility to embed scientific explanation into a more general frame, that of scientific inquiry. This move should enhance our insight into scientific practice and help us bridge the gap between the philosophical reconstruction and its object, i.e. scientific knowledge, gap that made older accounts appear artificially and alienated from real science. We’ll avoid this way the philosopher’s temptation to capture the ‘grand’ sense of scientific explanation that was questioned by some critiques of the philosophical accounts of explanation and understanding.\(^{12}\)

\(^{10}\) She directs her critique against the semantic conception of scientific theories.

\(^{11}\) Nancy Cartwright, “Models and the Limits of Theory: Quantum Hamiltonians and the BCS Model of Superconductivity,” in Models as Mediators. Perspectives on Natural and Social Science, ed. Mary Morgan et al. (Cambridge: Cambridge University Press, 1999), 244.

\(^{12}\) Such a critique was voiced for example by J.D. Trout in “Scientific Explanation and the Sense of Understanding,” Philosophy of Science 69 (2002): 212-233, against the philosophical approaches on scientific understanding as Salmon’s or Kitcher’s ones.
III. Models would fit the bill

In order to implement such a local, dynamic and non-theory-driven approach my suggestion is that by considering models as the bearers of explanations we would situate us in the right direction. Models could provide the adequate frame for a local approach on explanation, since they are usually built on local considerations. Such an approach will not be a theory-driven one if we consider the recent positions in the philosophy of science seeking to restore the importance and autonomy of models. Such authors as Nancy Cartwright, Margaret Morrison or Mary Morgan\(^\text{13}\) have shown that an adequate view on models and modeling activity could not be gained through a theory-centered type of conception.

Models can also in a good sense be viewed in a dynamic perspective and enhance this way a dynamic approach on explanation. For a general perspective, a dynamic approach seems to be proper for constructs that were ‘traditionally’ claimed to do heuristic work.\(^\text{14}\) As Weinert argued in his paper on *Models, Theories and Constraints*,\(^\text{15}\) models in comparison to theories exhibit fewer constraints. Therefore from a practical view a dynamical perspective on models should be more manageable than one on theories.

Last but not less important is the fact that a modeling approach could avoid the view on explanation as an application. Explanation as a construction process could be reflected either through the process of model generation or the one of model manipulation. The modeling view provides us therefore a concrete frame in which we could investigate the process of building an explanation.

Besides the above characterization, we might also bring some general additional clues to bear on the above choice. One such reason comes from the fact that the debate on explanation proved that pragmatics has to play an important role. Bas van Fraassen is the author that articulated in the eighties a pure pragmatic account on explanation. Nevertheless his position was criticized as being too unconstrained.\(^\text{16}\) On one side, models as scientific units in comparison to other sorts of constructs incorporate a very important pragmatic component. It is

---

\(^{13}\) Some important papers are gathered in the volume *Models as Mediators. Perspectives on Natural and Social Science*, ed. Mary S. Morgan and Margaret Morrison (Cambridge: Cambridge University Press, 1999).

\(^{14}\) In the “received view” models were mainly seen as heuristic means, dispensable after the formulation of the laws in the new domain.


Explanation Through Scientific Models: Reframing The Explanation Topic

this characteristic that should direct our attention to models as explanatory providers. On the other side, they would provide the necessary constraints for a pragmatic approach on explanation, constraints that could to be found by investigating their structure, building and functioning.

Another quite general reason for looking after explanatory virtues of models comes from the recent developments in the modelist camp. One of the central ideas of the recent modelist approach was to emphasize the mediator role that models have. Models are scientific constructs that mediate between theories, and phenomena. This fact qualifies them even better for explanatory jobs since explanation involves precisely this type of effort, namely, of seeking to get the theoretical constructs to bear upon phenomena.

***

Now taking a look into the debate on explanation we could notice that neither of the major accounts made reference to models as serious candidates for the role of explanation bearers. In “Aspects of Scientific Explanation” Hempel dedicates an entire section to the discussion of explanation through models. He arrives at the conclusion that models cannot offer genuine explanations due to their intrinsic limitations as scientific units. His position actually unfolds in the frame of “received view” towards models, view that conceives models as being scientific constructs of secondary importance for scientific knowledge. One might find it curious that the inhibition towards considering models explanatory outlived the dismissal of Hempel’s account during the debate on scientific explanation.

Though we would not find philosophical accounts that explicitly tried to address explanation through models, there are authors that called for such an enterprise or touched on the issue while pursuing different philosophical aims. Such is Cartwright’s thesis in her 1983 book How the Laws of Physics Lie in which she states her simulacrum account on explanation rather as a manifesto that remained much unsubstantiated. For Cartwright “to explain a phenomenon is to find a model that fits into the basic framework of the theory and that thus allows us to derive analogues for the messy and complicated phenomenological laws

17 This idea is mainly articulated in Margaret Morrison and Mary S. Morgan, “Models as Mediating Instruments,” in Models as Mediators, 10-37.
which are true of it."\(^{19}\) Besides this programmatic claims, her position is rather offensive and critical towards the ‘classical’ view on laws and Hempel’s conception of explanation: “The simulacrum account is not a formal account. It says that we lay out a model, and within the model we ‘derive’ various laws, which match more or less well with bits of phenomenological behaviour. But even inside the model, derivation is not what the D-N account would have it be.”\(^{20}\) Her position seemed to be isolated in the philosophical landscape of that time. Only in the ‘90s we find few more accounts taking into consideration explanation through models. Such is Hughes\(^ {21}\) account that addresses a particular kind of explanation, the one through theoretical models. His proposal is part of a collective effort of more philosophers especially in the second part of the ‘90s (such as N. Cartwright, M. Morrison, Mary Morgan) that aim to restore the importance of the modeling topic in philosophy of science. Their view backed a rather more pragmatic and pluralistic approach on models. In Morrison pragmatic view models are autonomous agents in the production and manipulation of scientific knowledge. Though her investigations reveal new aspects of the modeling activity it doesn’t say much about explanation through models. Under the general claim that “it is models rather than abstract theory that represent and explain the behavior of physical systems”\(^ {22}\) she adds that “a model explains the behavior of the system because it contextualizes the laws in a concrete way.”\(^ {23}\) But this “concrete way” cannot be made more explicit. This could be seen as a limitation of her pure pragmatic and functionalist approach.\(^ {24}\) In fact we could say more generally that the main consequence of the work of the modelistic camp in the ‘90s was to reveal the richness of the issues related to the modeling activity and to stimulate the research on the topic.

---


\(^{22}\) Morrison, “Models as Autonomous Agents,” 53.

\(^{23}\) Morrison, “Models as Autonomous Agents,” 64.

IV. A plausible general frame for approaching explanation through scientific models

I’ll further present a sketchy account on explanation through models advanced by Roman Frigg and Stephan Hartmann (the LOOP account as they call it). Though it was never published but only presented at a conference it lays out in my opinion a promising direction to be followed. After presenting it I’ll point to some further necessary steps that I see as immediate entries in a working agenda seeking to implement this approach.

Frigg and Hartmann’s account unfolds under a representational approach on models – approach that sees models as primarily being representations – but according to the authors it does not require an explicit account of representation. They claim to follow the previous suggestions of Rom Harré and Nancy Cartwright (already discussed). For Harré and Cartwright explanation provides us a picture of the facts and it is given through representation in models. So in their account the representational function of models is central. Nevertheless, they don’t require a worked out theory of representation in order to articulate their account. Only a general picture is assumed through which model represents a target system or some part of it.

In their account the explanandum is a feature or propriety of the target system or an event or phenomenon within this system. They exclude other types of explanadum from their account. The explanans on the other side is the model itself. The problem is then: how does a model M explains an occurrence O exhibited by the target system T that is represented by M? Their account specifies four steps that make out the process of an explanation. The first two steps are called identification steps. In the first one we identify the occurrence in the target (OIT as they call it), i.e., the behavior of interest in the target system that has to be explained. Using their example, Boltzmann ideal gas model in which the gas is represented through an ensemble of a huge number of hard balls moving in a confined space under Newtonian classical laws, the OIT is the expansion of the gas in the entire volume of the container when a separating wall is removed. In the second identifying step, the occurrence in the model (OIM) is identified, i.e., the element in the model that corresponds to the occurrence in the target that we wish to explain. In our example it corresponds to the spreading of the balls in the entire volume.

The next two steps are called the explanatory steps. In the first one, called explanation_1 we have to reproduce the OIM in the model, meaning that the OIM has to follow from the basic assumptions of the model. ‘Follow’ is not made more explicit in any way but is not reducible to deduction as in Hempel’s model. In the mentioned example of the ideal gas one has to show that the approach to equilibrium follows from the assumptions about the balls (the fact that are hard, that they collide elastically, etc). The fourth and last step, called explanation_2, involves the translation of knowledge obtained in the model (and about the model) to the target system. In our example, we know that the balls bounce around such that they reach the equilibrium distribution (what Boltzmann proved) and that the balls are idealizations of a certain kind of the molecules. This way what holds true in the model approximately carries over to the real system.

Having laid out the above sketchy account before proceeding further I want to emphasize two important points. The first one is the fact that the above approach does not have to be taken as a general schema for all sorts of scientific explanations. I do back up the conviction that is shared by many philosophers today that the variety of scientific explanations cannot be captured by a single general schema. The LOOP approach addresses a specific type of explanation – the one through scientific models – and it is possible that even this scope has to be better qualified. The second point I want to make is that the LOOP schema is an empty, quite unsubstantiated schema that has to be filled out. The authors recognize this and see the needed content to be delivered by the different types of representations corresponding to different explanatory strategies. I will address this issue in subsequent paragraphs. Nevertheless both above observations emphasize the need of articulating a solution in the sense of a “local philosophy of science,” as Nick Huggett calls them. Huggett characterizes such a “local philosophy” as a position in which “philosophical problems are to be found and treated using the resources of more-or-less delineable scientific programs” and “not by trying to make science fit some prior vision.”

A last additional observation is related to the precaution the authors take regarding the issue of explanation and truth. I completely agree with their position that an explanation is an explanation due to its ‘inner constitution’ and not for how good, bad or fruitful it is. We have to bite the bullet if we are going to consider explanation through scientific models.

---

Now, I think that there are some urgent tasks to be addressed in order to be able to begin filling out the schema. This shows also the directions that I think we have to move. The first one is related to the authors’ specification about the different types of representations. I think this is a particular way of approaching the issue of contextualizing the schema and it is too connected to the problem of scientific representation. We make this way the solution dependent on a solution to another philosophical problem that proved to be recalcitrant. Instead of looking too far we should take some baby steps and engage in exploratory investigations. In this sense we might first look at different types of models and the way they are used in providing explanations. A causal mechanism, for example, in molecular biology, will constitute a different sort of explanation than a theoretical model exhibited through a differential equation. So, the further investigation should take into account specific fields and subfields of science in order to fill out the schema.

An immediate task that we have to undertake if we are to search for filling the schema is to get more clarification on the four steps invoked by the schema. The least problematic seems to be the third step, i.e. the explanation or explanation in the model. All the other steps raise issues that had to be addressed in the first move. As in the previous paragraph, I think that a proper clarification can be done only in the limits of a specific scientific context, comprising a specific type of explanation in a particular scientific field.

The two identification steps are not unrelated and the ‘correspondence’ between the two occurrences, i.e., OIT and OIM, has to be clarified in the larger frame of the ‘correspondence’ or the relation between models and the target system. We have to take into account the fact that the two identification steps engage different sorts of means: the OIT involves rather an experimental setting (as in the ideal gas example) meanwhile OIM uses conceptual means. Nevertheless the account could cover (though this is not intended by the authors) also cases in which we model the target system through another material system – as would be the case of a scale model.

The last step appears as the most unusual one. Nevertheless I see it as a nice feature of the schema due to the fact that it did not end the explanatory act in its final product but opens it to the potential implications it might have. It facilitates this way the embedding of explanatory inquiry into the more comprising frame of scientific inquiry. The step is defined by the transfer of knowledge obtained in the model to the target system. This action should relate us to another important issue, which I see as necessary to be addressed. It is the issue of understanding.

In the literature of scientific explanation there are positions that ignore or even reject as inappropriate any approach on understanding (as Hempel’s one); but
there are also approaches that seek to account for understanding. Friedman manifestly requires that an explanation account should provide also an insight on how understanding is gained and how it relates to explanation. For Friedman scientific understanding is a ‘global affair’ given by the global aspects of the explanation, i.e., “the relation of the phenomenon in question to the total set of the accepted phenomena.” Following Friedman’s ideas, scientific understanding is construed through the connection to the unification of a body of knowledge in the unificationist accounts of Kitcher, Schurz or Barthelbort. These positions assume the global character of understanding and are also the most articulated accounts on understanding.

In order to claim understanding from models we’ll have to look at the local level and the specific characteristics that a local sort of understanding might have. Salmon makes room for such a sort of understanding (besides the global one) – understanding of causal mechanisms – in the frame of its causal account of explanation. In the case of models we should expect a variety of types of understanding corresponding to the different types of representations they exhibit.

There are fundamental differences from the sort of global theoretical understanding that was usually considered. We should rather focus on understanding in practice (as Morrison also briefly suggested). In this frame a quite important difference is the one between the understanding obtained in the model and the one claimed over the target system. Hi analyzed this distinction for theoretical models from condensed physics but we might consider it for other types of models. The last step of LOOP-schema extends the understanding gained in the model to the system. This sort of understanding has to be qualified by the specific epistemic engagements that the modeling action involves. The understanding that builds up at this moment could be claimed as the one delivered through models. One might claim also the existence of the other moments of understanding that are generated in the other steps of the schema. In this case we need to know how they related to each other and what is the impact on the final understanding.

So, overall, we might say that there is some urgent work to do in order to make the LOOP-schema a workable solution. The main message is that the

30 Salmon, Causality and Explanation, 79-92.
schema has to be implemented on specific contexts and it might provide a fruitful
general frame for investigating the explanatory virtues of models of specific kinds.

Conclusions

In order to conclude we might say that the actual landscape of philosophy of
science and the recent results from the last two decades require the
reconsideration of the classical topic of scientific explanation. Instead of
dismissing it as irrelevant one should take seriously the challenges of resettling it
in a new frame of inquiry. I’ve tried to suggest such a frame and to argue for its
pertinence. My argumentation draws on the recent advancements in the
philosophy of science on the subject of scientific models that would provide this
general frame. It reveals also the fact that we might encounter a plethora of new
sorts of problems that we have to address. Nevertheless as it stands now a new and
rich field of research opens for the interested philosopher of science.