How Scientific Models Can Explain

Alisa Bokulich
Boston University
abokulic@bu.edu

Abstract:

Scientific models invariably involve some degree of idealization, abstraction, or fictionalization of their target system. Nonetheless, I argue that there are circumstances under which such false models can offer genuine scientific explanations. After reviewing three different proposals in the literature for how models can explain, I shall introduce a more general account of what I call *model explanations*, which specify the conditions under which models can be counted as explanatory. I shall illustrate this new framework by applying it to the case of Bohr's model of the atom, and conclude by drawing some distinctions between phenomenological models, explanatory models, and fictional models.

I. Introduction

Models are used in a variety of ways in scientific practice; they can, for example, function as proto-theories, pedagogical devices, or as tools for generating and testing hypotheses.[1] In what follows I shall defend the view that—in some cases—models can perform an *explanatory* function as well. Despite the widespread use of models to explain phenomena in science, as well as the vast philosophical literature on models and explanation individually, remarkably little has been said in the philosophical literature about the

[1] William Wimsatt (1987) provides a nice taxonomy of the various ways in which scientific models can be false, and the various functions that such false models can serve in scientific research. Regrettably he does not talk about the explanatory function of false models, which is my focus here.

explanatory function of models. I shall begin by briefly reviewing three different proposals in the literature for how models can explain. These proposals are presented in what I think is an increasing order of their strength. The first is Carl Craver's (2006) account of "mechanistic model explanations, " the second, Mehmet Elgin and Elliott Sober's (2002) "covering-law model explanations," and third Ernan McMullin's (1978; 1985) "causal model explanations." Abstracting from these three proposals, I shall give a general account of what it is that makes an explanation a *model explanation*, and articulate the conditions under which it is reasonable to take such models as being genuinely explanatory. Using this general framework I shall introduce a fourth type of model explanation that is more typically found in fundamental physics. I shall briefly illustrate this framework using as an example a familiar model in physics that I shall argue is explanatory despite involving a fictional representation. Finally I shall conclude by drawing some important distinctions between "fictional models", "phenomenological models" and "explanatory models."

II. Three Approaches to Model Explanations

While there is a growing recognition that models have a legitimate role to play in scientific explanations, there remains considerable disagreement over precisely how and when models are to be counted as explanatory. Recently Carl Craver has defended the view that models can be explanatory in the context of the neurosciences. According to Craver, "models are explanatory when they describe mechanisms. Perhaps not all explanations are mechanistic. In many cases, however, the distinction between explanatory and non-explanatory models is that the latter [sic], and not the former [sic], describe mechanisms" (Craver 2006, p. 367).[2] Mechanistic explanations are a kind of constitutive explanation, in which the behavior of a whole is explained in terms of the operation and interaction of the mechanism's parts. His version of model explanations is then best described as *mechanistic model explanations*.

Craver places rather stringent conditions on when a mechanistic model is to count as genuinely explanatory. He writes, "In order to explain a phenomenon, it is insufficient merely to characterize the phenomenon and to describe the behavior of some underlying mechanism. It is required in addition that the components described in the model should correspond to components in the mechanism in [the target system] T" (Craver 2006, p. 361). Craver elaborates on this requirement by drawing a distinction between "how possibly models" and "how actually models." A how possibly model, he explains, merely describes how a set of parts and activities might be organized such that they produce the explanandum phenomenon (p. 361). How actually models, by contrast, describe the "real components, activities, and organizational features of the mechanism that in fact produce the phenomenon" (p. 361). In other words, for a mechanistic model to count as genuinely explanatory, it must correctly reproduce the *actual* mechanisms. More specifically Craver requires that "a

[2] Presumably Craver means "the former, and not the latter".

mechanistic explanation must begin with an accurate and complete characterization of the phenomenon to be explained" (p. 368). In my view, this requirement for a model to be explanatory is far too strong. If one has a complete and accurate description of the phenomenon, it is not clear to me that one has a model at all. Indeed this sounds much closer to a theoretical description of the system, than a model. I think that RIG Hughes was absolutely right to say that "To have a model . . . is *not* to have a literally true account of the process or entity in question" (Hughes 1990, p. 71). Hence Craver's account succeeds in defending the view that models can explain, only by reducing the notion of a model to a complete and accurate description of the system.

A second, and in my view slightly stronger defense of the view that models can explain has been given by Mehmet Elgin and Elliott Sober. They are concerned specifically with a class of models in evolutionary biology known as "optimality models"; these models describe the value of a trait that maximizes fitness, given a certain set of constraints. They argue that despite the idealizations involved in these optimality models, they are nonetheless genuinely explanatory. They write,

> Optimality models contain idealizations; they describe the evolutionary trajectories of populations that are infinitely large in which reproduction is asexual with offspring always resembling their parents, etc. . . . We want to argue that optimality models are explanatory despite the fact that they contain idealizations. (Elgin and Sober 2002, p. 447)

The sort of model explanations that Elgin and Sober consider are a species of covering-law causal explanations. In order to qualify as a *covering-law* causal explanation, the explanans must, first, describe the cause (or causes) of the explanandum; second, the explanans must cite a law of nature; and third, all of the explanans propositions must be true.[3] The further condition usually added to a covering-law explanation, namely that the explanans explains the explanandum by entailing it or conferring probability on it, is not satisfied by covering law model explanations, they argue, because it does not make sense to talk about the probability of an idealized circumstance (the one described in the model) that does not obtain in the real world.[4] Note that on Elgin and Sober's account of model explanations, the explanans must both cite a law of nature and be entirely true. For this reason I shall refer to their account of model explanation as "*covering-law model explanations*."[5]

Elgin and Sober also recognize the need for distinguishing those model explanations that can be understood as being genuinely explanatory from those

[3] It is unclear how Elgin and Sober can impose this last condition (of all the explanans propositions being true) insofar as the example they give does involve the false conditions they describe.
[4] Sober (personal communication) has subsequently clarified that the difficulty is specifically with conditionalizing on an idealization that involves an *impossible* situation (such as an infinite population size). Even if one grants that the probability of an event conditionalized on an impossible state of affairs has a well-defined value, it still posses a problem for the inferential link between explanans and explanandum.
[5] Note that "covering law model explanation" should be distinguished from "covering law model *of* explanation;" while in the latter 'model' just means 'an account', in the former 'model' is meant to convey the fact that the explanans makes essential reference to a scientific model.

that are not. In a manner similar to Ernan McMullin (whose views will be discussed below), their account proceeds by way of a de-idealization analysis. They write, "The idealizations in a causal model are *harmless* if correcting them wouldn't make much difference in the predicted value of the effect variable. Harmless idealizations can be explanatory" (Elgin and Sober 2002, p. 447). That is, on their view, a model can be explanatory only so long as the idealizations employed in the model are harmless.

Unlike Craver's complete and accurate descriptions, the models Elgin and Sober describe do involve idealizations, and hence are legitimately understood as models—not theoretical descriptions. However one can still criticize Elgin and Sober for giving an inadequate defense of model explanations insofar as the idealizations they cite do not seem to play any real role in the explanation itself—that is, the point of the harmless analysis is to show that the idealizations do not get in the way of those true parts of the model that do the actual explanatory work. Indeed their requirement that the explanans be entirely true seems once again to rule out the possibility that the idealizations or fictionalizations in the model can *themselves* be explanatory. A more robust defense of the view that models can explain should be able to describe how the idealizations *themselves* are capable of doing some real explanatory work.

One of the strongest defenses of the view that models can explain is also one of the earliest, and that is Ernan McMullin's work on what he calls "hypothetico-structural" (or HS) explanations and the status of idealizations in science. In a hypothetico-structural explanation, one explains the properties of a complex entity by postulating an underlying structural model, whose features are causally responsible for those properties to be explained. He notes that HS explanations are usually metaphorical and tentative, and that a good structural model will contain the resources to lay out an extended research program for the further refinement of the model.

McMullin explicitly distinguishes this type of explanation from nomothetic explanations, noting that although HS explanations may involve laws of nature, "the explanatory character of the model comes not just from the laws governing the constituents of the model but also from the structure in which these entities are combined" (pp. 146-147). For McMullin, these model explanations are ultimately to be understood as a species of causal explanations: He writes, "such explanations are causal, since the structure invoked to explain can also be called the cause of the feature being explained" (McMullin 1978, p. 139). For this reason I think McMullin's account is best labeled as *causal model explanations*.[6]

In this 1978 paper, McMullin does not address the question of how it is that a hypothetical and metaphorical model can legitimately be said to explain at all. For his answer to this question we must look to another paper of his on

[6] Although this label regrettably does not capture McMullin's emphasis on structures, I would nonetheless like to reserve the term "structural explanation" for a distinctive type of explanation that is neither nomothetic nor causal, as will be discussed below.

Galilean idealization, where he writes, "Our interest here lies in the use of models as idealizations of complex real-world situations, the 'falsity' this may introduce into the analysis, and the ways in which this 'falsity' may be allowed for and even taken advantage of" (McMullin 1985, p. 258). He continues,

> Every theoretical model idealizes, simplifies to some extent, the actual structure of the explanandum object(s). It leaves out of [the] account features deemed not to be relevant to the explanatory task at hand. Complicated features of the real object(s) are deliberately simplified in order to make theoretical laws easier to infer, in order to get the process of explanation under way. (McMullin 1985, p. 258)

McMullin then goes on to provide a taxonomy of two important ways in which scientific models can involve idealizations. He notes that on the one hand, models may idealize by simply leaving unspecified those features of the system that are deemed irrelevant. This is perhaps the less interesting notion of idealization. On the other hand, one may also need to idealize in the sense of simplifying or even omitting altogether those features of the system or object that *are* known to be relevant to the particular explanation being offered. He argues that it is important to distinguish these two types of idealization, for they both play an important role in the subsequent process of legitimating the models and showing that they are genuinely explanatory.

For McMullin, the justification of a model proceeds through a process that he calls *de-idealization*. That is, one goes through a process of "adding back in" those features that were omitted, or de-simplifying those assumptions that were included, but only included in over-simplified way. It is through this (sometimes lengthy) process of de-idealizing, that McMullin sees metaphoric models as laying out a research program. He notes that for such a process to work, the original model must successfully capture the real structure of the object of interest. When this process of de-idealizing can be given a theoretical justification, then McMullin concludes that the model can be counted as genuinely explanatory. When, on the other hand,

> techniques for which no theoretical justification can be given have to be utilized to correct a formal idealization, this is taken to count against the explanatory propriety of that idealization. The model itself in such a case is suspect, no matter how good the predictive results it may produce. (McMullin 1985, p. 261)

McMullin's approach, then, can be summarized by the following three claims: First, that idealized models can be genuinely explanatory; second, when they do explain, what they offer is a species of causal explanation (namely the structures postulated in the model are taken to *cause* the relevant observed features of system to be explained), and third, we are justified in taking the model as genuinely explanatory when a theoretical justification can be given for each step in the de-idealization process, that is, the de-idealization does not simply involve an ad hoc fitting of the model to empirical data.[7]

---

[7] McMullin is well known for taking this account of model explanation one step further, and arguing that when, in this process of de-idealizing a model, new discoveries are made and even more experimental data can be accounted for, then we have strong (though not conclusive) evidence for the existence of the structures postulated by the model (McMullin 1985, p. 262; and especially McMullin 1984). This further argument will not be discussed here.

Although McMullin's account of model explanations provides the strongest defense of the three approaches described here, his account is still subject to a number of limitations. First, it is not clear that all idealizations can be subject to a de-idealization analysis.[8] Hence a broader justificatory framework will be required to handle explanatory models that involve idealizations of this sort. Second, not all models represent their target systems via an idealization—some scientific models represent via a fictionalization. Ideally a more complete account of model explanations would at least countenance the possibility of fictional models explaining. And finally, McMullin's account is limited to specifically causal explanations. Although many explanations are causal, there are domains of science, such as quantum mechanics, in which many of the relevant explanations cannot be thought of as causal.

In what follows I would like to suggest a new, more general account of model explanations with the following features: First, my account of model explanation contains the previous three types of model explanations as special cases. Second, it can be applied not only to idealized models but fictionalized models as well. And third, it provides a framework for seeing how, in some cases, the fictionalizations and idealizations themselves can do real explanatory work.

III. A General Account of Model Explanations

[8] For a defense of this claim see Batterman (2005)

Despite the wide variety of models used in scientific practice—and the various explanatory purposes to which they can be put—I believe that there is nonetheless a set of core features that can be used to define a type of explanation that I have here called 'model explanation.' First, and perhaps most importantly, what makes something a *model* explanation is that the explanans in question makes essential reference to a scientific model, and that scientific model (as I believe is the case with all models) involves a certain degree of idealization and/or fictionalization. Second, a general characterization of model explanations requires giving an account of what it is in virtue of which, that these models can be said to be genuinely explanatory. My answer here draws on a suggestion made by Margaret Morrison , who writes, "The reason models are explanatory is that in representing these systems, they exhibit certain kinds of structural dependencies" (Morrison 1999, p. 63). Unfortunately, however, Morrison does not indicate how this suggestion is to be fleshed out into a philosophical account of scientific explanation, nor how it will be able to distinguish genuinely explanatory models from those phenomenological models that merely "save the phenomena."

There is, however, a similar account of scientific explanation that has been worked out in much greater detail by James Woodward. According to Woodward, an explanation can be understood as providing information about a pattern of counterfactual dependence between explanans and explanandum (Woodward 2003, p. 11). He fleshes out this idea of counterfactual dependence in terms of what he calls "what-if-things-had-been-different questions," or "w-questions" for short. That is, "the explanation must enable us to see what sort of difference it would have made for the explanandum if the factors cited in the explanans had been different in various possible ways" (Woodward 2003, p. 11).

While I think that Woodward's account is largely right, where I wish to part company with his view is in his construal of this counterfactual dependence along strictly manipulationist or interventionist lines. Woodward takes one of the key features of this pattern of dependence to be that it, in principle, permits one to *intervene* in the system in various ways. It is precisely this manipulationist construal that restricts Woodward's account of scientific explanation to specifically *causal* explanations.[9] As I shall argue in more detail below, I think it is a mistake to construe all scientific explanation as a species of causal explanation, and more to the point here, it is certainly not the case that all model explanations should be understood as causal explanations. Thus, while I shall adopt Woodward's account of explanation as the exhibiting of a pattern of counterfactual dependence, I will not construe this dependence narrowly in terms of the possible causal manipulations of the system.

So far I have articulated two key features of a general account of model explanations: first, the explanans must make essential reference to a scientific model, and second, that model explains the explanandum by showing how the

[9] Woodward in facts admits that perhaps not all scientific explanations are causal explanations, and allows that his theory may be extended in the way I am suggesting.

elements of the model correctly capture the pattern of counterfactual dependence of the target system. More precisely, in order for a model M to explain a given phenomenon P, we require that the counterfactual structure of M be isomorphic in the relevant respects to the counterfactual structure of P. That is, the elements of the model can, in a very loose sense, be said to "reproduce" the relevant features of the explanandum phenomenon. Furthermore, as the counterfactual condition implies, the model should also be able to give information about how the target system would behave, if the elements described in the model were changed in various ways.

In addition to these two requirements, a third condition that an adequate model explanation must satisfy, is that there must be, what I call, a further "justificatory step." Very broadly, we can understand this justificatory step as specifying what the domain of applicability of the model is, and showing that the phenomenon in the real world to be explained falls within that domain. This justificatory step is intended to call explicit attention to the detailed empirical or theoretical process of demonstrating the domain of applicability of the model. In other words, it involves showing that it is a good model, able to adequately capture the relevant features of the world (where 'relevant' is determined by which questions the model is specifically trying to answer).[10]

[10] This justificatory step does not simply involve showing that the model is empirically adequate. As I will discuss in more detail below, it is to be understood as playing a role analogous to Hempel's condition of truth (which he refers to as the "empirical condition of adequacy"), in so far as the justificatory

Although the details of this justificatory step will depend on the details of the particular model explanation in question, there are typically two general ways in which such a justification can proceed. In the first case, the justification might proceed "top down" from theory; that is, one might have an overarching theory that specifies what the domain of applicability of the model is—where and to what extent the model can be trusted as an adequate representation of the world.

More frequently, however, there is no such overarching justificatory theory, and the justification must instead proceed from the ground up, through various empirical investigations. We can see the importance of this latter sort of justificatory step in the examples due to McMullin and Elgin and Sober, in which the justificatory step took the form of a de-idealization analysis.[11] As I noted earlier, however, such a bottom-up de-idealization procedure will typically only work when the target system is related to the model system "smoothly" via an idealization. For those models that represent their target systems via some relation other than idealization, such an approach will not typically work. On my view, this justificatory step is important in so far as it plays a central role in distinguishing those models that are merely phenomenological, "saving the phenomena," from those models that are genuinely explanatory.

---

step is intended to rule out as explanatory those models that we know to be merely phenomenological.

[11] Even more typically (and perhaps even to a certain extent in all these cases) there is a combination of top-down and bottom-up justificatory procedures being employed together.

With this general characterization of model explanations in hand, we are now in a position to recognize that there are in fact several different "subspecies" of model explanations. Indeed so far we have encountered three different species of models explanations: Craver's *mechanistic* model explanations, Elgin and Sober's *covering law* model explanation, and McMullin's *causal* model explanations. This suggests that even after one has identified an explanation as a model explanation, there still remains the question of what *kind* of model explanation it is. Answering the question of what kind of model explanation it is, requires articulating what we might call the "origin" of this counterfactual dependence. For example, is the dependence of the explanandum on the elements cited in the model to be construed as either, the elements represented in the model *causally producing* the explanandum (in the case of causal model explanations), the elements of the model *being the mechanistic parts which make up* the explanandum-system whole (in the case of mechanistic model explanations), or the explanandum being a *consequence of the laws* cited in the model (in the case of covering law model explanations).

In addition to these three types of model explanations, there is an important fourth type of explanation, which has been largely overlooked in the literature, and this is what I call a *structural model explanation*. Following Peter Railton (1980), R.I.G. Hughes (1989), and Rob Clifton (1998), a structural explanation is a distinctive type of scientific explanation that is neither causal nor nomothetic. Very broadly, a structural explanation can be understood as one in which the explanandum is explained by showing how the (typically mathematical) structure of the theory itself limits what sorts of objects, properties, states, or behaviors are admissible within the framework of that theory, and then showing that the explanandum is in fact a consequence of that structure.[12] Structural explanations are typically found in fundamental physics, where causal explanations may be inappropriate. Examples of structural explanations include the explanation of the Pauli exclusion principle (Railton (1980), the explanation of the invariance of the speed of light in special relativity (Hughes 1989), and the explanation of correlations in the vacuum state of a relativistic quantum field (Clifton 1998).

Extending this account of scientific explanation to models, one can define a *structural model explanation* as one in which, not only does the explanandum exhibit a pattern of dependence on the elements of the model cited in the explanans, but in addition, this dependence is a consequence of the structural features of the theory (or theories) employed in the model. As we shall see next, structural model explanations provide one possible framework for making sense

---

[12] This definition of structural explanation is my own, and is not exactly identical to the definitions given by other defenders of structural explanations, such as Railton, Hughes, and Clifton. Nonetheless I think this definition better describes the concrete examples of structural explanations that these philosophers give, and is preferable given the notion of 'model' being used here.

of how fictionalized models—at least in some instances—can be genuinely explanatory.[13]

IV. Example of an Explanatory Fictional Model: Bohr's Atom

It is helpful to briefly illustrate this view that fictional models can be explanatory by applying the theoretical framework I have just outlined to a concrete example. A well-known example of a scientific model is Niels Bohr's model of the hydrogen atom. According to Bohr's model, the electron can orbit the nucleus only in a discrete series of allowed classical trajectories known as stationary states. While in a stationary state the energy of the electron is constant, and the electron can only gain or lose energy by jumping from one stationary state to another. When such a transition or "quantum jump" occurs, a single photon of a given frequency is emitted (or absorbed). The frequency of the photon is given by the difference in energy of the two allowed orbits. The spectrum of hydrogen (such as the Balmer series shown at the bottom of Fig. 1) is built up out of the photons being emitted in these jumps between stationary states, where only those frequencies (or wavelengths) corresponding to allowed quantum jumps occur, and the intensity (or brightness) of a spectral line is given by the probability of that jump occurring. So, for example, the red spectral line

[13] Obviously not all models are explanatory. For example, one would not say that Descartes's vortex model or Ptolemy's epicyclic model actually explain the motion of the planets. It would be a mistake, however, to conclude from this that *no* models are explanatory. The conditions under which we can consider a model to be genuinely explanatory are given in this paper.

H_α, is the result of the electron jumping from the n=3 orbit to n=2 orbit, and the green spectral line H_β is the result of transitions from the n=4 to n=2 orbit.

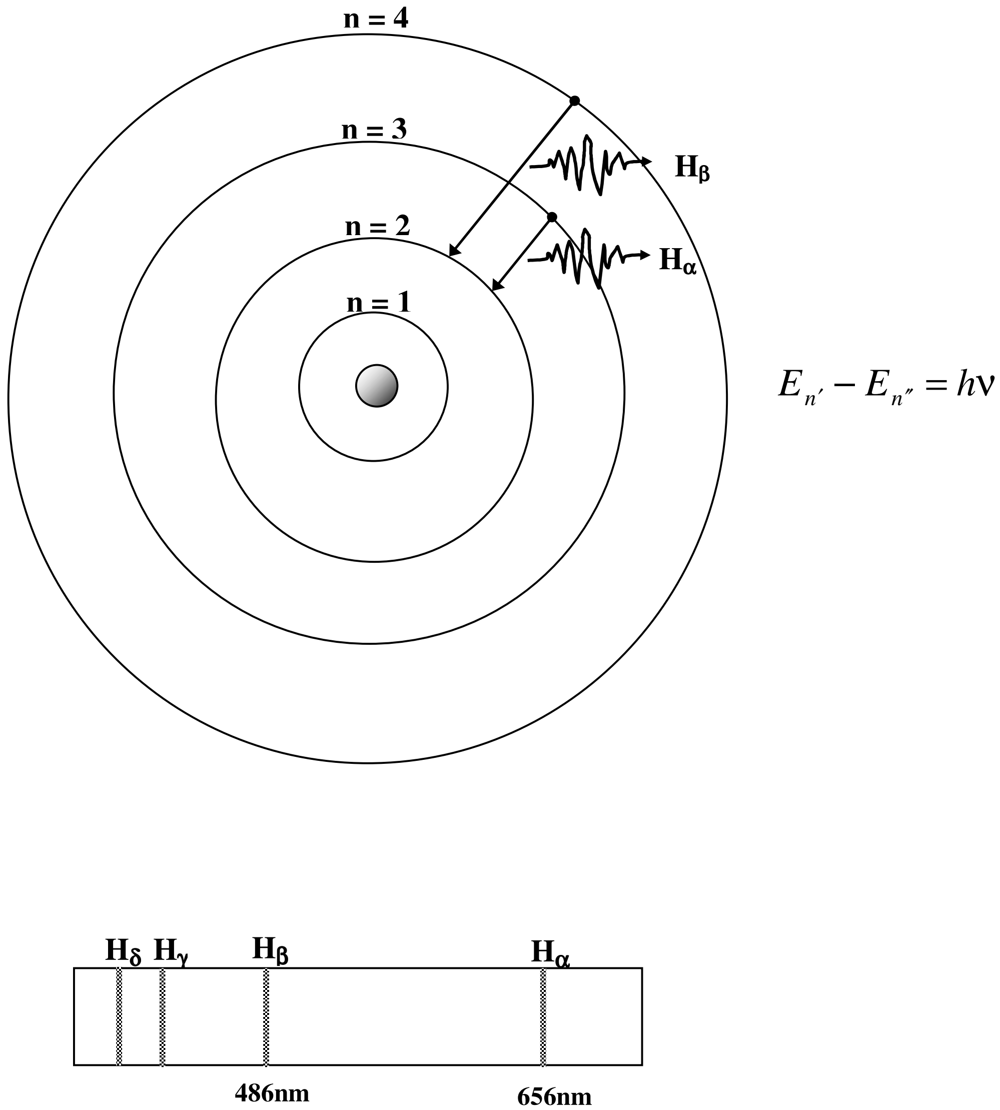

Figure 1: (*Top*) Bohr's model of the hydrogen atom. The nucleus is in the center and the *n* label the electron orbits in the stationary states. When the electron jumps from a higher stationary state to a lower one, it emits a photon (indicated by squiggly arrow) whose frequency (and hence also wavelength, $\lambda = c/\nu$) are given by the difference of energy, *E*, of the two stationary states in accordance with the equation given. (*Bottom*) The Balmer series of the emission spectrum of hydrogen. The H_α spectral line, for example, is produced by photons emitted by

the electron jumping from the n=3 to n=2 stationary state and has a wavelength of 656 nanometers.

From his model, Bohr was able to theoretically derive the Balmer formula

$$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \qquad n = 3,4,5\ldots\,, \qquad (1)$$

which Johann Balmer had constructed by an ad hoc trial and error directly from the empirical data, as well as the more general Rydberg formula. Moreover, Bohr was able to derive an expression for the Rydberg constant R

$$R = \frac{2\pi^2 me^4}{h^3 c} \qquad (2)$$

in terms of more fundamental constants of nature (Planck's constant h, the speed of light c, and the mass m and charge e of the electron) —and this theoretical value for R was in precise agreement with the empirically determined value. Bohr's model also explained the apparent capriciousness of the spectral lines, that is, why only certain jumps between stationary states were allowed, and this explanation was based specifically on an analysis of the classical motion of the electron in a stationary state orbit.[14]

As we know well today, however, these Bohr orbits are *fictions*—according to modern quantum mechanics the electron in an atom does not follow a definite classical trajectory in a stationary state and is instead better described as a cloud of probability density around the nucleus. I want to defend the view

[14] Bohr's selection rule, explaining why only certain transitions between stationary states occur, is given by his well-known—though widely misunderstood—correspondence principle. For a discussion of Bohr's correspondence principle see Bokulich (2008a), Chapter 4.

that despite being a fiction, Bohr's model of the atom does in fact explain the spectrum of hydrogen. McMullin, for example, writes of Bohr's model of the atom,

> The fact that the Bohr model worked out so remarkably indicates that the structure it postulated for the H atom had some sort of approximate basis in the real. . . . Later QM would modify this simple model in all sorts of fundamental ways. But a careful consideration of the *history* of the model . . . strongly suggests that the guidance it gave to theoretical research in quantum mechanics for an immensely fruitful fifteen years must ultimately have derived from a "fit" of some sort, however complex and however loose it may have been, between the model and the structure of the real it so successfully explained. (McMullin 1968, 396)

Now it is somewhat surprising that McMullin grants Bohr's model this explanatory status, since I do not think that the Bohr orbits can be properly thought of as an "idealization" of the true quantum dynamics. Moreover, one cannot recover the modern quantum-mechanical description of the hydrogen atom from the Bohr model through a de-idealization analysis—that is, one cannot simply add something back in to the Bohr orbits that was left out or change some parameter of the orbit to recover the correct quantum description. These two pictures of the atom are simply not related continuously to each other in this way. Hence, I think Bohr's model is best described as a fictional—not idealized—description of the hydrogen atom.

Despite being a fictional representation, I do think McMullin is right to say that Bohr's model genuinely explains the spectrum of hydrogen. We can understand how this fictional model explains by means of the general philosophical framework of "model explanations" that I outlined earlier. Recall that in a model explanation the following three conditions hold: First, the explanans makes reference to an idealized or fictional model; Second, that model explains the explanandum by showing that the counterfactual structure of the model is isomorphic (in the relevant respects) to the counterfactual structure of the phenomenon.[15] This means that the model is able to answer a wide range of "what-if-things-had-been-different" questions. And third, there is a justificatory step specifying what the domain of applicability of the model is and that the model is an adequate guide to that domain of phenomena.

Applying this framework, then, to the example of the Bohr atom, we can say that the counterfactual structure of Bohr's model is isomorphic to the counterfactual structure of the spectral phenomena. That is, there is a pattern of counterfactual dependence of the emission spectrum of hydrogen on the elements represented in Bohr's model—specifically the motion of the electron in the discrete stationary states and the quantum jumps between them. Furthermore, Bohr's model is able to correctly answer a number of "what-if-things-had-been-different questions," such as how the spectrum would change if the orbits were elliptical rather than circular, or how the spectral lines would change if the hydrogen atom were placed in an external electric field (the Stark effect), and so on. This shows that Bohr's model is not simply an ad hoc fitting of the model to the empirical data, as would be the case in a merely

[15] The notion of isomorphism is being used loosely here.

phenomenological model. Finally, modern semiclassical mechanics provides a top-down justificatory step showing that Bohr's model—despite failing as a literal description—is nonetheless a legitimate guide to quantum phenomena in certain domains.[16]

One might object that the Balmer series is only truly explained by modern quantum mechanics, and that the explanatory power of Bohr's model is merely illusory. Although this objection is intuitively plausible, I think it is misguided for at least two reasons. First, one might ask: Why stop at modern quantum mechanics? Why not require that the only true explanation be the one offered by the more fundamental quantum field theory? Or even the long-sought-after more complete theory of quantum gravity? Second, and more generally, I think it is a mistake to believe that there can only be a single legitimate scientific explanation for a given phenomenon. A closer examination of scientific practice reveals that, not only can there be more than one genuine scientific explanation for a given phenomenon (what might be called the *explanatory pluralism thesis*), but that some of these explanations may turn out to be deeper than others.[17] To say that one scientific explanation is not as deep as another is not the same thing as saying that it is no explanation at all. Hence on my view, Bohr's model does genuinely explain the Balmer series, though the explanation it offers may not be

[16] For an introduction to modern semiclassical mechanics see Bokulich (2008a), Chapter 5 and references contained therein.

[17] For a nice discussion of the notion of explanatory depth see Hitchcock and Woodward (2003).

as deep as that offered by modern quantum mechanics, and moreover, the explanation offered by modern (nonrelativistic) quantum mechanics may not be as deep as that offered by quantum field theory.[18]

I noted earlier that even after one has identified an explanation as a legitimate model explanation, there still remains the question of what *kind* of model explanation it is. In the case of Bohr's model of the atom, I do not think that it is properly thought of as a causal model of the phenomena. Instead it seems closer to what I described as a structural model explanation—that is, one in which the explanandum can be seen as a consequence of the structure of the theory or theories employed in the model (in this case a particular blending of classical and quantum mechanics), which limits what sorts of objects, properties, states or behaviors are admissible within the framework of that theory.

I would like to conclude by drawing some distinctions between fictional models, phenomenological models, and explanatory models. A *phenomenological model* is only of instrumental value to the scientist. Often—though not exclusively—they are constructed via an ad hoc fitting of the model to the empirical data. Phenomenological models are useful for making predictions, but they do not purport to give us any genuine insight into the way the world is. An *explanatory model*, by contrast, does aim to give genuine insight into the way the

[18] Although I have here made it sound like the notion of explanatory depth always parallels the hierarchy of fundamental theories, elsewhere I have defended the view that, in some cases, the less fundamental theory may provide a deeper explanation than the more fundamental theory. See Bokulich (2008b).

world is. Finally, by *fictional model* I mean simply a model that represents the world by means of fictional entities, states, or processes (and to distinguish fictional models from idealized models—specifically fictional entities or processes that are not related to the true ones in the world by what might be thought of as a distortion or series of successive cases).[19]

There is a general assumption that all fictional models must be merely phenomenological models—and I think this is mistaken. I have tried to argue that fictional models can, in some cases, nonetheless give us genuine insight into the way the world is. Specifically, they can do so by correctly capturing in their fictional representations, real patterns of structural dependencies in the world.

[19] For example, frictionless planes are better thought of as an idealization rather than a fiction, because they are conceptually related to actual planes by a limit—a series of successive planes in which the friction is made to be less and less.

Works Cited

Batterman, R. (2005), "Critical Phenomena and Breaking Drops: Infinite Idealizations in Physics," *Studies in History and Philosophy of Modern Physics,* 36: 225-244.

Bokulich, A. (2008a), *Reexamining the Quantum-Classical Relation: Beyond Reductionism and Pluralism*. Cambridge: Cambridge University Press.

Bokulich, A. (2008b), "Can Classical Structures Explain Quantum Phenomena?", *British Journal for the Philosophy of Science* 59 (2): 217-235.

Clifton, R. (1998) "Scientific explanation in quantum theory", Unpublished manuscript on PhilSci Archive: http://philsci-archive.pitt.edu/archive/00000091/.

Craver, C. (2006), "When mechanistic models explain", *Synthese* 153: 355-376.

Elgin, M. and E. Sober (2002), "Cartwright on explanation and idealization", *Erkenntnis* 57: 441-450.

Hitchcock, C. and J. Woodward (2003), "Explanatory generalizations, part II: Plumbing explanatory depth", *Noûs* 37: 181-199.

Hughes, R.I.G. (1989), "Bell's Theorem, ideology, and structural explanation", *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem.* Notre Dame, IN: University of Notre Dame Press.

Hughes, R.I.G. (1990), "The Bohr atom, models, and realism", *Philosophical Topics* 18: 71-84.

McMullin, E. (1968), "What Do Physical Models Tell Us?", in B. van Rootselaar and J. Staal (eds.) *Proceedings of the Third International Congress for Logic, Methodology and Philosophy of Science* (Amsterdam 1967). Amsterdam: North Holland Publishing Co: 385-396.

McMullin, E. (1978), "Structural explanation", *American Philosophical Quarterly* 15: 139-147.

McMullin, E. (1984), "A case for scientific realism", in Jarrett Leplin (ed.) *Scientific Realism*. Berkeley, CA: University of California Press, pp. 8-40.

McMullin, E. (1985), "Galilean idealization", *Studies in History and Philosophy of Science* 16: 247-273.

Morrison, M. (1999), "Models as autonomous agents", in M. Morgan and M. Morrison (eds.), *Models as Mediators: Perspectives on Natural and Social Science.* Cambridge: Cambridge University Press: 38-65.

Railton, P. (1980), *Explaining Explanation: A Realist Account of Scientific Explanation and Understanding.* Ph.D. Dissertation, Princeton University.

Wimsatt, W. W. (1987), "False models as means to truer theories", in M. Nitecki and A. Hoffman (eds.) *Neutral Models in Biology*. Oxford: Oxford University Press.

Woodward, J. (2003), *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.