May 2022

New Thinking About Models

Aaron Alexander Kruk
University of Wisconsin-Milwaukee

NEW THINKING ABOUT MODELS

by

Aaron Kruk

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

in Philosophy

at

The University of Wisconsin-Milwaukee
May 2022

ABSTRACT

NEW THINKING ABOUT MODELS

by

Aaron Kruk

The University of Wisconsin-Milwaukee, 2022

Under the Supervision of Professor Joshua Spencer

Contemporary philosophers of science have been wholly concerned with understanding models through their ability to represent their target systems. According to these 'representationalists' understanding how models represent will answer the foremost philosophical questions pertaining to scientific models. I propose a new way to think about models. I argue that two of the functions that models preform, explanation and exploration of their target systems, are codependent on one another. That is, a model is capable of explanation if, only if, and because it is capable of exploration (and vice versa). From this codependency, it follows that we need not (and cannot) understand these two functions—and, *a fortiori*, fully understand models *simpliciter*—in terms of representation. I conclude by outlining a new research program pertaining to models that asks about these non-representational functions and analyze them in such terms.

Table of Contents

0 Introduction

Recent tradition in philosophy of science has focused on the role of scientific models (henceforth 'models') as one of the primary vehicles of scientific knowledge. Inside of this tradition, many philosophers have treated questions of how models represent the world—or, more specifically, their target systems—as the foremost regarding their study. Giere (1988; 2004) argues that models represent in virtue of being similar to their target systems and seeks to analyze models as such. Structuralists, like van Fraassen (1980; 2008), Bueno (1997; 2010), French (2003; Bueno and French 2011), and Ladyman (Bueno, French, and Ladyman 2002), argue that models represent by being isomorphic (or partially isomorphic) with their target systems. Inferentialist approaches, defended by the likes of Suárez (2003; 2004; 2015), assert that a model represents a system if we can use the former to make inferences about the latter.[1] Callender and Cohen (2006) spearhead the General Gricean camp which contests that models represent target systems by an agent's action of stipulative fiat. And Frigg (2010a; 2010b), Nguyen (Frigg and Nguyen 2016; 2021), Godfrey-Smith (2006), and Salis (2021) all defend the fiction view of models where a model represents a system in virtue of being part of some sort of fictional discourse or game of make-believe.

Regardless of the differences between these accounts, each agrees that we should understand models in terms of representation—that is, how they refer to their targets, how we can believe that the entities in the models stand for or are entities in the target system, and how these representations can be more or less accurate. According to the 'representationalist' camp, answering these questions and their derivatives purports to give us the majority of, if not a complete, philosophical picture of scientific models.[2]

In this paper, I propose to look at models from a much different perspective. Among other functions, scientific models often allow an agent to explain features of a target system and further

[1] Suárez' version of inferentialism is perhaps most closely related to what I argue for here. In one sense, I agree with him that we should deflate some of the questions regarding representation and models. However, he and I differ in that I do not reject that there are interesting questions about how models represent that outstrip answers to questions about their non-representational capacity. Likewise, I think that representation is a separable function of certain inferential functions that models serve (i.e., exploratory and explanatory functions); Suárez does not.

[2] I've been intentionally brief with the nuances of the above accounts in order to get straight to my own points. For a more detailed overview of models, representation, and these competing accounts, see Frigg and Nguyen 2020.

explore its other features. I argue that there is a vital connection in the way that models accomplish these two functions; namely, models are able to explain in virtue of their ability to explore and they are able to explore in virtue of their ability to explain. This codependency is best seen by looking at the role that models play in experimentation. Briefly, models help scientists to prescribe and conceive of experiments which are essential drivers of the development and articulation of scientific theory. Recognizing this facet of experimentation is a strong step towards recognizing the codependency of these two functions in models.

Building off of this claim, I ultimately suggest that the tradition of primarily and wholly thinking about models in terms of representation is misguided. The overarching thesis of the paper is thus both revisionary and prescriptive. If we are to understand how models work, we should redirect a good measure of our philosophizing about them in terms of representation to other functions which must be understood non-representationally; including, but not limited to, explanation and exploration.

The paper will proceed as follows. In §1, I begin with some preliminaries about explanation and present my own account of scientific exploration. To illustrate how these functions work in models, I consider the history of the evolution of the atomic model from Thomson to Rutherford which I will use throughout the paper. In §2, I present my argument for the codependency of explanatory and exploratory functions of models. I then provide a prescriptive method for generalizing my conclusion to any model. In §3, I build on the claim established in §2 to argue for my overall thesis. First, I show that the explanatory and exploratory functions involved in this codependency are separable from a model's representational functions. That is, the successful performance of explanation, exploration, and their codependency in a model is not dependent whatsoever on a model's ability to represent. I then demonstrate how it follows from this and the claim established in §2 that to fully and primarily understand models we must do so, at least partially, in non-representational terms. I then conclude by sketching a new research program which must be pursued for any philosophy of scientific models.

1 Preliminaries

1.1 Explanation in models

It is broadly accepted in the literature that at least some models function to explain features of their target systems (Woodward 2003; Strevens 2004; 2008; Cartwright 1983; Elgin and Sober 2002) and some function to explore other features (Fisher 2006; Gelfert 2016, ch. 4; Massimi 2018; 2019).[3] Beginning from this fact, I will argue that there is a codependency between these two functions. Before doing so, however, some preliminaries regarding explanation and exploration are needed. None of what I argue here is particularly reliant on any complete account of scientific explanation (of which there are many) or scientific exploration (of which there are few, if any). For the former, all that is required for my argument to go through are oft agreed upon claims that are relatively invariant between competing accounts. For the latter, I present some of my own conservative claims to characterize it, which are at least plausible, if not uncontroversial.

Beginning with explanation, there is a long history going back to Hempel's Deductive-Nomological model of explanation in which philosophers have generally agreed that scientific explanation requires at least three features. They must be (a) putatively veridical (Hempel (1965; Strevens 2008), (b) provide answers to why-questions (van Fraassen 1980; Bromberger 1966), and (c) be supported or motivated by empirical evidence (Hempel 1965).[4,5] These are not the only more generally agreed upon features of explanation. Others include it being a member of an interrelated family of modal concepts such as 'laws of nature', 'counterfactuals', and 'causation' (Hempel 1965).[6] While such extra features are interesting in their own right and may even promise intriguing ques-

[3]I note here that later on, in §2.2, that the functions which these authors recognize as exploratory also serve an interesting explanatory function which I dub *secondary explanation*.

[4]There are some exceptions to these general agreements. Most notably, Bas van Fraassen (1980) argues for a wholly pragmatic account of explanation where an explanans need not be veridical. What I argue here intends to target the majority of views and, accordingly, will circumvent discussions of the pragmatic view. I will note, however, that some modifications to my account could, in principle, allow it to be co-opted by the pragmatist or instrumentalist more generally. I will return to this idea in §3.2.

[5]We should not conclude that since explanations must be supported or motivated by empirical evidence, they cannot be motivated theoretically. I entirely accept that they can and often are. However, I do assume that this is not a necessary feature of scientific explanations.

[6]Again, there are exceptions to this general agreement as well. See Woodward 2003.

tions in the context of how models explore and explain, they are not at work in what I argue here.

Explanations figure into models in a unique way. A model embeds an explanans as a structure which has its explanandum in the target system (an explanandum may often be an observation or measurement made on that target system), among other structures in the model. This allows that there be different structures in any model which do not function as explanans for the same explanandum, and which might have entirely non-explanatory functions as well (e.g., structures that have exploratory functions). The only requirement these different structures need have is that they, and their functions, are consistent with those of the embedded explanans. (Briefly, I assume a model consists of structures which perform different functions—e.g., explanation, exploration—for the model. Each structure is, in principle, capable of performing multiple functions in a single model. As to what metaphysical or ontological categories these structures fall into, I remain agnostic towards here as I do not think it has any import in my arguments. In general, I avoid metaphysical assumptions throughout the paper).

The claim that models embed explanans as structures follows immediately from the facts that the function of scientific explanation can be realized without a model and, when realized in a model, it must be accompanied by other structures. As to the former claim, anyone who accepts that there are alternate vehicles of scientific knowledge (e.g., theories) that do not reduce down to models will concede this. Theory may also explain parts of the world, independent of any model.

However, this is not universally accepted; some reject that theories do not just reduce to collections of scientific models (Cartwright 1983). For this contingent of philosophers, the claim that explanations are realized elsewhere than in models can still be accepted on the grounds that an explanation can be given as a response to a why-question without reference (implicit or explicit) to any particular scientific model. So long as this answer is empirically motivated and putatively veridical, it will count as a scientific explanation. This also leads us to my second claim, that a single explanans in a model must be accompanied by other structures. Answering a why-question may be sufficient for an explanation, but it is not sufficient for making a model. Other structures must compliment the explanans to make it a proper model. Thus, explanans are embedded as structures in models.

This argument can be illustrated by introducing the running example I use throughout the paper: the progression of the atomic model from Thomson to Rutherford.

Two of the key structures of the Thomson (1904) model of the atom were that it posited localized negative charges inside of it (electrons) which were set in a homogeneous, positively charged medium. The postulation of the former was mostly motivated by the results of Thomson's cathode ray experiments where rays (beams of electrons) where deflected when run through a magnetic field inside of a Crookes Tube. The latter structure was partially inherited from Dalton's atomic theory but also empirically motivated by a range of experiments including the establishment of ions (partially from work done by Faraday 1833). Thomson posited it to account for the net neutral charge of the atom, since there must have been some opposite positive charge to balance out the measured negative charge of the electron structure.

These two structures eventually received the nicknames of the "plum" and "pudding" structures, respectively; the electrons similar to plums embedded in a homogeneous medium of pudding in plum pudding. The former structure served as the explanans for the results of Thomson's cathode ray experiments; the existence of charged subatomic particles would explain why the cathode ray was deflected in experimentation. The latter functioned in part as an explanans for the belief that atoms had no net electric charge; although it would also be somewhat theoretically motivated and play an exploratory role, which we'll see in the following subsection.

The Rutherford (1911) model, which superseded Thomson's, retained the electron structure. In lieu of the 'pudding' structure, however, it posited a centralized positive charge in the atom which made up the majority of the atomic mass: the nucleus structure. This was postulated in response to results of experiments performed by Rutherford's post-doctoral assistants, Geiger and Marsden. In their eponymous "Geiger-Marsden" experiment, the researchers bombarded a sheet of gold-foil with a beam of positively charged alpha particles to observe the rates of deflections at different angles. The rates predicted by the Thomson model were far from those actually observed; a statistically significant amount of the particles were deflected at rates greater than 90°.

The electron structure in the Rutherford model still served the explanatory work that it did in the Thomson model, and the nucleus structure still did the explanatory work that the pudding structure did, but the latter also did more explanation. It accounted for the results of the Geiger-Marsden experiment where the pudding structure did not. The observed deflection rates of incident alpha particles on gold atoms would be explained by the nucleus structure, but not the pudding.

While these structures in the models functioned to explain phenomena in the target system (the atom), the models themselves were not needed for any explanation to be realized or independently posited. Thomson could claim, as the answer to the why-question "Why did I observe the results of the cathode ray experiment that I did" that negatively charged subatomic particles existed. This explanation is putatively veridical and empirically motivated, but realized linguistically without any need for reference to the model. This point is further driven home by the fact that Thomson postulated his model *after* the positing of electrons as an explanans. It follows that explanations need not be realized by a model.

So, a model embeds the explanans of an explanation which has its explanandum in the target system of the model. This explanation may be realized independent of the model (either in a theory or linguistically) and only properly becomes part of a model when complimented with other structures. It follows that a model is not exhausted by a single embedded explanans and the explanatory function it performs; it will have other structures which partake in other functions. One of these functions, I argue in the following subsection, necessarily needs to be exploration.

1.2 A theory of exploration

As to scientific exploration, much less focus in the philosophical literature has been poured into it than scientific explanation has had. Accordingly, there is little consensus as to what any account of exploration must include. To makeup for this gap in the literature, I propose that we fill in some of its features by positioning it in the 'empty spaces', so to speak, left between two other fields of study in philosophy of science: scientific discovery and scientific progress. Drawing from these areas, I propose my own conservative characterization of what scientific explorations essentially are.

Scientific discovery has most often been discussed inside the philosophical literature against the backdrop of Reichenbach's (1938) 'context distinction' between the 'context of discovery' (the *de facto* thinking processes used by scientists) and the 'context of justification' (the *de jure* correctness of these thoughts) for scientific inquiry. Reichenbach, among other 20th century philosophers of science, was convinced that only the context of justification was a proper object of study for philosophy of science; the context of discovery was left for the sociologists and psychologists, but had no import on *how* science should be done. This idea is perhaps most famously expressed by Popper (1959). According to him, "conceiving or inventing of a theory, seems to me neither to call for logical analysis not to be susceptible of it" (ibid.). The philosophy of science, and its quest to define a logic of scientific inquiry, consists solely in investigating the methods through which we test theories, not introduce them.

This distinction led to a dominant approach in 20th century philosophy of science to regard scientific inquiry as properly dividing into two parts: (i) the generation of new theories, ideas, and scientific knowledge and (ii) their validation. However, from the late 20th century into the 21st, a different approach to philosophy of science which focuses on the actual practice of science has advocated for a tripartite division of scientific inquiry which adds a division concerning the pursuit and articulation of new theory to the other two.[7]

The latter approach can actually be traced back to Whewell's (1840) work on the philosophy of discovery in the 19th century, predating Reichenbach and Popper. Whewell's work clearly separates inquiry into three elements: the non-analyzable 'happy thought' or eureka moment when a theory is first introduced; the process of 'colligation' or the clarification and explication of new theories; and the verification of the colligated theory.[8]

The reintroduction of this second division once again opened up the possibility for there to be

[7]That is not to say that Popper, Reichenbach, and their colleagues were completely oblivious to the process of theory development. For them, however, this was mostly lumped in to the context of discovery which did not, according to their beliefs, warrant philosophical study.

[8]Interestingly, Whewell was a polymath that also actively practiced as a physicist. That he did hands on work inside of science and was able to recognize this tripartite distinction in scientific inquiry lends credence to the fact that the 'practice turn' in contemporary philosophy of science and the reintroduction of this division is well founded.

a philosophy of discovery. While most all agreed that a logical analysis for the 'happy thought', eureka moments that correspond to the first division of scientific inquiry was impossible, if part of discovery includes aspects of this second division—the colligation of theory—then it will be apt for philosophical analysis.[9]

In analyzing this colligation division of inquiry, Hanson (1960) has argued that principles of abduction govern how we should develop competing theories in the face of anomalous new phenomena. Other approaches advocate for an analysis of rational, albeit non-formal, general scientific methodologies. Laudan (1980) sees these methodologies as both a guide for *how* to acquire novel scientific knowledge and *why* such knowledge is validated. Nickles (1984; 1985; 1989) has argued that the methods through which theories are developed and articulated also serve important functions for their justification. Insofar as justification consists in successful novel predictions derived from a theory, their justification relies on our ability to generate new knowledge. Schaffner (1993) dubs this generation as a "weak" evaluation procedure which provide reasons for accepting a hypothesis as promising and worthy of further attention, but does not evaluate it on standards of truth or falsity.

I observe a general thread that holds between these different accounts of how the articulation and development of new theory works: through the introduction novelty. This is not the sort of novelty which traditionally characterized scientific discovery, a la a new theory apprehended in one of Whewell's 'happy thoughts', but the advent of novel empirical evidence and ways to gather such evidence that go into developing an infantile fragment of theory into a more complete, testable one. This gain of novelty in our colligation process is what I propose as the central element of scientific exploration. Exploration is the function that produces the novelty that develops and articulates theory.

Exploration, in this sense, can be realized in a variety of ways; new data can be collected from previously untested experimental setups; new theories can allow us new ways to conceptualize old

[9]There are some contemporary philosophers, most notably Paul Thagard (1984; 1999; 2010; 2012), who argue that the eureka moment *can* be analyzed, either logically or in non-logical, yet rational terms. I'm inclined to agree with this, particularly when looking at neuroscience and cognitive science as extensions to philosophy, but only need to focus on the colligation division to make my point here. See also Magnani (2009).

observations; we may be directed to look at parts of the world that we haven't previously seen or studied; and exploration, in some cases, is the making of tentative predictions about what the world might be like or what might be observed that are instrumental in making new observations.

We can see this concept active in scientific progress as well. This is easily illustrated when looking at two competing accounts of progress between Kuhn and Popper. While there is far more literature on progress beyond these two, that their accounts are polar opposites of one another yet both essentially rely on my concept of exploration should lend weight to the idea that the concept is operative in any account of progress.

Popper (1963) holds that science progresses through experimental refutation of conjectured theories. Kuhn (1970), in direct opposition to Popper, says science progresses through the accumulation of insoluble puzzles for a (global) paradigm to the point of crisis until a new paradigm (at first local, then global) is produced that resolves these puzzles and does away with the old paradigm. On either account, there is a need for the sort of novelty that exploration provides. Popper requires that we have some way of producing refutations for our theories, canonically experiment. While we might be tempted to say this is just a part of theory validation, our third division of scientific inquiry, Popper grants (ibid.) that a single failed prediction does not falsify a theory. This is because conjectured theories often need to be articulated and developed both theoretically *and* experimentally for them to be the proper objects of refutation.[10] Kuhn, on the other hand, needs there to be anomalies that fail to conform to a paradigm for progress to occur. Such anomalies of course require novelty, in that they have not been previously assimilated into or accounted for by the prevailing paradigm, which typically comes as a result of surprising experimental results. Again, this is just what I've asserted explorations provide us with.[11]

Looking back to our running example of the atomic model, the pudding structure in the Thomson model primarily served an exploratory function for Rutherford and company. The postulation

[10]We might also argue here that scientific exploration plays an essential role in validation *itself*. Producing novel evidence is part of how we refute theories. This seems to be in line with what Nickles (1984; 1985; 1989) has argued. I'm inclined to agree with this point but will save further comments for §3.2. Interestingly, if this is the case, a study of exploration as a key part in both colligation and validation may wind up blurring the line between the two.

[11]Exploration will also play a key role in the development of a global paradigm from a local one.

of the structure allowed Rutherford to derive new experiments and predictions that would allow him, Geiger, and Marsden to inquire into different subatomic features of the atom.[12] This eventually resulted in new empirical observations, the development of the atomic theory, and, in some sense, contribution to scientific progress.

So, both progress and discovery require a sort of novelty provided for us by scientific exploration. This should give us a working, if incomplete, characterization of exploration for our purposes.

2 The codependency of explanation and exploration

2.1 Exploration in virtue of explanation

Models explain and explore in at least the senses illustrated in §1. I now argue that they are only able to perform either of these functions in virtue of performing the other. I propose this as the biconditional claim that models are able to explain if, only if, and because they are able to explore.[13] Conversely, they are able to explore if, only if, and because they are able to explain.[14]

I begin by establishing the biconditional's right-to-left direction. If a model is able to explore, this is because it is able to explain (and must do so). The argument for this conditional is rather simple. For any model to successfully explore its target system, it needs to rely on some prior knowledge of that system. Exploratory functions are able to target a system because of structures in a model that provide explanatory functions.

Before seeing how explanations do this, the point needs to be made that exploratory functions cannot explore a target system without being guided by some prior knowledge of the system. The

[12] Of course, these predictions and the ability to design their gold-foil experiment were also dependent on the electron structure, which will be a major point in the following section.

[13] Of course, there may exist a model that neither explains nor explores its target system, thus falling out of the scope of my biconditional. While I am not immediately aware of any such model, I may be inclined, if presented with one, to conclude there is a categorical difference between it and those which explain and explore. One could then understand what I argue here as analysis of the latter category.

[14] "Because" is not a symmetric operator so, strictly speaking, this latter statement is not the converse of the former. However, for expediency, I will talk as if it is symmetric and refer to the first of these two claims as "the biconditional".

accrual of novel information used to flesh out a fledgling theory requires some fragment of theory already *in hand* to be developed, along with a sort of ignorance pertaining to parts of the target system where the theory has not been developed. If exploratory functions were supposed to be interpreted as telling us what is already in the system or how it behaves (purporting to provide non-modalized *knowledge*) they would not be exploratory. Their purpose is to help develop the incomplete parts of a theory or model. In this sense, they require some knowledge to build off of, but cannot be interpreted or understood as giving us a literally true description of what the world is like. They are explicitly introduced without knowledge of whether they actually target anything, but in order to buffet ready had knowledge.

The knowledge that allows us to isolate a system for exploration is only provided by an explanation. Initially, we might be tempted to think it does so through representation: by providing us with a literally true description of what the system is like. This would be an appeal to the putative veridical nature of explanations. However, I argue that there is a separate feature of explanation which allows successful exploration of a target systems: by telling us how to build experiments to test a model.

Of course, some of these details in prescribing experimental setups will come from exploratory functions, but it crucially rests on explanation as well. One of the central features of scientific explanations noted in the previous subsection were that they are motivated by empirical evidence. Effectively, this motivation is what allows for explanations to tell us how to do the sort of empirical manipulations and interventions that give us interesting and telling results in the first place. That they are posited in response to empirical evidence from past experimental successes allows them to act as a foothold for conceiving new ways of experimenting on that system.

An anecdote from Popper (1963) helps demonstrate this. He once asked a group of physics students in Vienna to "Take pencil and paper; carefully observe, and write down what you have observed!". They were confused and unable to complete the task because the instruction, 'Observe!' is absurd. Observation is always selective. It needs a chosen object, a definite task, an interest, a point of view, a problem. In some cases, observation can be exploration which will need to presuppose one of the items just listed for it to be possible. Novelty in our colligation stage of inquiry cannot begin from nowhere. Explanations are the sort of chosen objects, tasks, interests, and points of view which allow for us to make novel observations in the first place.

To illustrate this, we can look back at the exploratory work that the Thomson model enabled Rutherford and his post-doctoral assistants to do that would result in them positing their new model. Thomson's model allowed for Rutherford, Geiger, and Marsden, to explore the atom's structure by designing new experiments (the Geiger-Marsden gold foil experiment), producing novel information (the observed deflection rates of incident alpha particles in the experiment), and revealing previously unknown aspects of the target system (the nucleus). However, it would have been incapable of doing any of this if it did not also serve some explanatory function for other aspects of the atom. To focus on the electron structure, which primarily performed explanatory functions in the model, its positing was empirically motivated by the results of Thomson's cathode ray experiments. This told Thomson, Rutherford, and their contemporaries about the possibility of subatomic particles and structure. This was previously inconceivable under Dalton's atomic theory which held that atoms were indivisible and the smallest units of matter. The explanatory work electrons did for these past experiments was able to point Rutherford, Geiger, and Marsden into new directions for how to run new experiments and further articulate atomic theory.

Without these structures serving this function, Rutherford and company would have been incapable of exploring the atom in the ways that they did. They could now conceive of the world in new ways, as being open and testable beyond the veil of the atom's exterior. That significant experimental results could be had from probing matter with charged particles was established by Thomson and his electron structures. They told Rutherford where to start when looking at the atom which ultimately enabled him to conceive of the gold-foil experiment, attempting to use a prior established explanation in a new way (i.e., through bombarding matter with positively charged alpha particles).

The surprising deflection rates measured from the experiments eventually brought Rutherford to propose a new model of the atom that did away with the homogeneous, positively charged medium that the Thomson model posited (the 'pudding') and suggested a different structure to explain both the net neutrality of the atom and this new empirical data (the nucleus). Thus, the pudding structure served to help explore the subatomic properties of the atom, by providing the empirical novelty necessary for the advancement and development of atomic theory. But it only did so by working off the back of the explanatory functions of the atom's structure in the Thomson model.

2.2 Explanation in virtue of exploration

I now turn to establish the left-to-right direction of the biconditional. If a model is able to explain, this is because it is able to explore (and must do so).

My argument for this goes back to the idea that a model *embeds* an explanans. An explanation does not need to be realized in a model. It may be realized purely verbally or perhaps by theory independent of any model. However, these latter realizations of explanations are not the same as their instantiations inside of a model. Simply answering a why-question does not mean one has suddenly produced a scientific model. A single why-question is often too narrow in scope to exhaust the questions we can or seek to ask about a target system. A single explanans must be accompanied by other structures to actually form a model. But, when new structures are introduced alongside an explanans, new sorts of explanatory functions are produced. It is these explanations that necessarily rely on the possibility of exploration.

Complementing an explanans with other structures results in new predictions and new potential answers to why-questions that outstrip those made by the explanans per se (or by the other posited structures, whether they serve explanatory, exploratory, or other functions). We should think that these explanatory functions have all the same logical features that are had between the originally posited explanans, its explanandum, and the future predictions it makes, but it notably differs in a variety of informal ways. First, these explanations are posited before we've explicitly formed any why-questions. We jointly derive new predictions from our explanans *qua* model structure and the other structures of the model and these structures together tell us why we *would* observe such predictions, if in fact observed. But the questions come, at best, concurrently with the explanations, not before. Second, their empirical motivation is only indirectly had from the empirical motivation behind the independent structures which jointly produce the new explanation. Again, since they precede their posited explananda, the observation of these explananda cannot directly motivate them. Lastly, the sort of epistemic attitude properly held toward them is not full belief in their putative veridicality as it would be for our initially embedded explanans. At best, they give us tentative descriptions of what a system might be like or how it might behave which we can be optimistic about, albeit not fully confident.

For the sake of clarity, I suggest we refer to the sort of explanation canonically thought of in our discussion of explanation in §1.1 as *primary explanation* and this latter sort just described as *secondary explanation*. A rough summary of the distinction is as follows:

- *Primary Explanation*: The canonical sort of explanation that is directly empirically motivated; it answers an explicitly formed why-question that is usually posed before the explanans is posited; the proper sort of epistemic attitude taken towards it is full belief in its veridicality.
- *Secondary Explanation*: The sort of explanation that is most uniquely realized in models; is a tentative or modalized explanation of what possibly could explain something; it is indirectly motivated by empirical evidence; it answers a not-yet-formed why-question that is posed concurrently or after the positing of the explanans; the proper sort of epistemic attitude taken towards it is optimistic, but something less than full belief in its truth.

That models perform secondary explanation has already been acknowledged elsewhere in the literature. Gelfert (2016) points out that some models provide proofs-of-principle and suggest how-possibly explanations. Massimi (2018; 2019) recognizes that certain models provide us with modal knowledge of what a target system might be like. And Fisher (2006) sees models as acting instrumentally towards analyzing the further predictions and features that might be derived from a given theory. These are evidently secondary, not primary, explanatory functions. They make predictions about unobserved observables, provide potential explanations that we can be optimistic about, if not fully committed to, and they start from other parts of theory that are better established or may already serve primary explanatory functions. Strikingly, however, is that these philosophers recognize these features not as a different sort of explanation, but as exploration! This is rather telling to the point I'm trying to make. When models perform secondary explanatory functions, they are only able to do so because they beget exploration.

When we go to test secondary explanatory models through experimentation, the predictions they make (whether confirmed or defeated) allow us to gather new information about the target system. Notably, primary explanations can err and thus be unsuccessful, or in need of partial revision. But when the predictions of secondary explanations are experimentally defeated, we do not similarly mark them as failures. They can be successful insofar as they assist in theory articulation. Part of this is that these sorts of explanations help us generate *more* explanations.

Experiment plays a key role here. We pit our models against the experiments they prescribe us to design and run.[15] And, through experiment, we are able to glean novel information about a system. When such predictions are defeated we're able to revise the structures which would have explained them, if observed, or postulate entirely new structures that serve as new explanans. Alternatively, when predictions are confirmed, we can revise the sorts of attitudes we held towards such structures to have stronger epistemic commitments—eventually to the point of developing the secondary explanation into a function of primary explanation. Secondary explanation can only do the explanatory work it does in models by also allowing for the type of theory development that I have classified here as exploration. Thus, models can only explain in the idiosyncratic way that they do because they are capable of exploring.

To see this in action with the atomic model, the Geiger-Marsden experiment was a prime instance of the type of experiment that resulted from secondary explanation. The joint derivation of different new predictions (i.e., expected deflection rates of incident alpha particles on gold foil)

[15] Another note on experiments should be made here. Not only do these sorts of explanations have to bring about exploration when we run experiments related to them, but such experiments cannot even be prescribed by the model without exploratory functions. If all we have is a single explanans structure, we cannot use it to generate any experiments that do other than confirm or disconfirm those explanans. The sorts of experiments that are prescribed by these tentative explanations have to have some degree of novelty that goes beyond the evidence used to motivate the explanans of a primary explanation. That such novelty is needed, however, just means that there needs to be some exploratory function that our explanatory function is wrapped up with.

made from the electron and pudding structures of the Thomson model allowed for such an experiment to be conceived. These two structures would explain these deflection rates, if observed, and a tentative confidence could be placed in their veridicality. By way of this secondary explanatory function performed by the model, Rutherford and company were able to design their experiment and ultimately acquire the type of novel information necessary for the development and articulation of atomic theory. The defeat of the predictions made by these secondary explanations were not failures, at least not in the sense that the primary explanation performed by the electron structure would have been a failure if Thomson's results from the cathode ray experiment were not replicated elsewhere, but successes in the colligation period of discovery.

2.3 A method for generalization

So far, we've seen how this codependency works in a single historical instance: the development of the atomic model from Thomson to Rutherford. While a powerful example, to assume that this sort of analysis is applicable to each and every model would commit a fallacy of over-generalization. I do not intend to make this claim, I only introduce this as a proof-of-principle. However, the example suggests a prescriptive method for seeing how this works in any case.

Consider any model and its target system. While there may be multiple phenomena that have been observed in that system, select one of them. Locate the structure in the model that is posited to explain this phenomenon. (It will help to phrase the phenomenon in a why-question; why do the relevant scientists, the model makers, think we observe it? The answer will give us the putative explanans that should be embedded in the model).

Next, consider the experiments which have been used to test this model. If such an experiment has resulted in progress, new explanations, or discoveries, and further development of the model, it will be indicative of the explanatory-exploratory codependency in models. This will be evident in historical cases where a model has been superseded by a stronger model or in contemporary cases where an experiment prescribed by a model has illustrated some flaw in the model without yet having a replacement.[16]

Such experiments can be used to identify the key structures which perform secondary explanation and exploration in the model. So long as these structures were motivated independent of one another (empirically for the explanatory structure that, theoretically or instrumentally for the exploratory structure), were integral in designing the experiment identified, and jointly derived predictions that were tested in the experiment, we should have a clear case of the codependency of secondary explanation and exploration in the model. As long as both a secondary explanation function (i.e., a possible explanans for yet unobserved explananda) and an exploration function (i.e., a way to gather novelty for theory development) are identified together, we should have confirming evidence for this codependency.

Again, this is only a procedure. I am optimistic that it will show how the universality of the phenomenon I've presented so far in a robust range of models, but the analysis still needs to be done. Of course, it may fail in a sufficient amount of cases where a model is able to explore without secondary explanation or able to perform secondary explanation without being capable of exploring. If this be the case, my position will have been falsified. Even so, this would not be a total defeat. In the progress of philosophical theory development, falsification of a promising, albeit fledgling theory is as important as confirmation. At a meta-level, the testing of my model of models can do the sort of exploratory work that I assert all explanatory models do.

[16] A flaw in this procedure for generalization is that certain models may not have been used yet, or at all, to design and run experiments. If this is a contemporary case, further details may be needed. Have experiments not been run due to financial or technical limitations? Will they be run? If it is a historical case or a contemporary case where we do not expect experiments to be run using the model, there may be reasons to be inclined to treat it as categorically different than the sort of model I am interested in here. The real worry would be finding a model which performs secondary explanation or exploration without the other function, which I discuss in a moment.

3 Understanding models non-representationally

3.1 Explanation and exploration without representation

To remind the reader, I am ultimately attempting to argue that our philosophical approach towards scientific models should not be primarily (i.e., always the first approach taken) or wholly (i.e., the only approach taken) in terms of how models represent. I then conclude in the next subsection by sketching a new research program which must be pursued in the philosophy of scientific models. The argument for my overarching thesis is the present purpose of this subsection. It runs as follows:

The Anti-Representation Argument

P1. A significant portion of models can explain their target systems if, only if, and because they can explore those systems (and they can explore only because they can explain).

P2. Any model's (secondary) explanatory and exploratory functions, as well as the codependency of these functions, are not dependent on that model's representational capabilities. Accordingly, these former functions and their codependency cannot be understood in terms of representation.

P3. If a significant portion of models perform functions that cannot be understood in certain terms, then models *simpliciter* cannot be understood primarily or wholly in such terms.

C: Models cannot be understood primarily or wholly in terms of how they represent their target systems.

So far, I've argued that any model which performs what I've called *secondary* explanations can and must only due so because of its ability to perform exploratory functions (and vice versa). If my procedure for generalization introduced in §2.3 is applicable to a robust enough class of models, then we will have enough to establish the first premise of the argument.

While the first premise is crucial, it is not sufficient for establishing our conclusion, even if buffeted by a conditional similar to the third premise. Establishing the second premise (and incorporating it in the antecedent of third premise) avoids a vital issue that would otherwise defeat us from making an inference from the first premise to our conclusion. This problem, which I've not yet ruled out, is that explanation, exploration, and their codependency, may ultimately depend on a model's ability to represent. If this is the case, then we will likely have to understand these functions in terms of representation.

It will help to build up the potential objection I anticipate here before countering it in order to both situate my response as well as to avoid any straw men. One of the key features of primary explanations is that they are putatively veridical. In virtue of this, the explanans embedded in models will function as representations of aspects of the target system—that is, they should be understood as purporting to give a literally true description of what the target system is like where the explanans refers to something in the target system. Likewise, the proper epistemic attitude towards such structures is full belief.[17]

This is where the problem can be pressed. Even if there is a codependency between exploration and secondary explanation in models, since at least some of the structures which perform these functions must represent the target system when they perform *primary* explanations, we have prima facie reason to suspect that representation is somehow operative in any function that these structures perform. If exploration and secondary explanation are dependent on its operation, we will not be able to use either of these functions, or their codependency, in arguing that a complete understanding of models cannot come in terms of representation. The deepest understanding of how these functions work would bottom out in how they represent their target systems. Thus, we must establish the second premise to rule out this possibility. To do so, I argue in a piecemeal fashion; first establishing that exploration does not depend on representation, then that secondary explanation does not, and finally that neither does their codependency.

The first task is rather easy. A model's ability to explore cannot be dependent on representation because successful exploration does not need the structures which perform it to be veridical, i.e., to refer to anything in the target system. In fact, as we've seen with the explorations made possible by the pudding structure of the Thomson model, the failure of a structure to represent anything is often crucial in successful exploration for a model. To try to understand exploration in representational terms, then, would be a misnomer.

The argument that secondary explanation does not depend on its structure's ability to represent

[17]That is not to say explanations are just representations. True descriptions, for example, are not always proper answers to why-questions. If answering the question "Why did the glass shatter but not the mug when knocked off the table?" the description "The cat knocked them both over" may be true of the situation, but not answer the why-question and, subsequently, not provide an explanation.

something in the target system is a bit more nuanced. We may want to say that the function of secondary explanation is parasitic on primary explanation which, evidently, seems to be dependent on representation. To show that this is not the case and that secondary explanation is separable from representation, we should begin by looking back at our atomic models.

In Rutherford's use of the Thomson model to design his gold-foil experiment, the electron and pudding structures jointly cooperated to produce secondary explanatory functions, including predictions of the deflection angles of the incident alpha particles on the foil. While the electron structure functioned as a primary explanation for the results of Thomson's cathode ray experiments, this function played no role in the derivation of these new predictions. The effect that electrons would have in the Geiger-Marsden experiment was *neglected* and thus ignored in Rutherford, Geiger, and Marsden's calculations due to their relatively low weight compared to the incident alpha particles. While the model's electron structure still served as a primary explanation for these *past* experiments (and would be preserved in the superseding model) this function had no part in the secondary explanation enabled by the model.

The point that this example illustrates is that the sorts of new experiments used to test models in the process of theory development need not rely or use the primary explanatory functions of its structure. Generalizing now, primary explanation is *retroductive*, it works in response to what has been observed, what empirical evidence we have gathered, and what experiments have been run in the past. Secondary explanation, on the other hand, functions as essentially conjecture. It tells us how to build experiments, how to test our models, and plays an essential role in the process of scientific exploration. And, most notably, *it precedes primary explanation*. Effectively, the two sorts of explanation differ only as a matter of confirmation and proper epistemic attitudes held towards them. When a model functions for a scientist as a secondary explanation, it is part of what helps positively develop primary explanations.[18]

[18]That secondary explanation is not dependent on representation and that we cannot understand the former in terms of the latter is also indicated by the sort of attitudes that are proper to take towards structures which serve exploratory functions. Unlike primary explanations, they only require weaker epistemic attitudes (e.g., lesser degrees of confidence or weak belief) as to whether the exploratory structures provide an accurate description of the system, or whether they refer to actual entities in the target system. There may be rational requirements to adjust these attitudes after experimentation, but this will have been after the function of secondary explanation has done its work.

This flips the direction of dependency that the original objection which P2 of the Anti-Representation Argument sought to avoid. Secondary explanation does not depend on a model's ability to function as a primary explanation for anything; but a model's ability to function as a primary explanation depends on the past success of secondary explanation! Given this reversal of dependency, we no longer have reason to think secondary explanation need be understood in terms of representation.

The same points made for exploration's and secondary explanation's separability from a model's ability to represent can be co-opted for their codependency. Overall, that neither function is dependent on representation should act as a defeater toward any reasons for thinking that their codependency is likewise dependent on representation. But a further point can be made here. These two functions, and their codependency, are vitally tied to the process of experimentation. Experiments, in general, are not readily understood in terms of representation. They're direct manipulations performed on the target system, not representations of that target system. That models prescribe us to make such experiments completely foregoes questions of representation. Thus, we need not and cannot understand them representationally.

The majority of this section has worked towards establishing P2. Last but not least, some support needs to be given towards our third premise, the conditional of the argument. My reasoning for this is simple. This conditional is an instance of a more general inference schema about how to understand a population with heterogeneous sub-populations. We can only understand the whole in virtue of understanding its sub-populations. But, if there are two (or more) significant sub-populations which require different approaches and methodologies to understand them, and these approaches are not applicable to the other sub-population(s), then we'll need different ways to understand these groups to understand the whole.[19]

As a helpful analogy, biology and particle physics as two different sub-fields of science. Granted that they are both significant (although a particularly proud biologist or physicist might disagree), saying that we could use the tools of physics to understand all of science would be absurd. Biology,

[19] What counts as 'significant' for a population will of course be context-dependent as well as likely partially-dependent on the attitudes or interest of an inquiry. This is not immediately a problem for my account, since this will need to be hashed out in further study of the codependency I've recognized here as outlined in §2.3.

and its subject matter, cannot be understood with such tools. To understand science as a whole, we'd have to understand biology in its own terms (i.e., by looking at its instruments). Similarly, if we're to understand scientific models, we'll have to understand the class of models who's significance I've argued for here. Since we cannot understand these in terms of representation, we cannot understand models *simpliciter* wholly or primarily in such terms. We'll have to philosophize about models in non-representational terms at least partially and primarily (i.e., as a first approach to certain questions and kinds of models).

3.2 The sketch of a new program

If we cannot understand the class of scientific models that have this explanatory-exploratory codependency in terms of how they represent their target systems, how are we supposed to understand them? What sorts of questions are we to ask about them? What tools can we use to philosophically analyze scientific exploration and explanation? If what I've argued so far is correct, these questions, among others, require our attention. I conclude that paper by sketching brief answers to them here, which will point us to a new research program surrounding the philosophy of scientific models.

Before sketching how I think such a research program should go about, I'd like to avoid a potential misinterpretation about what I am arguing here. I am not advocating for a sort of holistic anti-realism or instrumentalism about scientific models. I myself am a realist and what I've argued here is entirely compatible with a robust realism about models. I fully concede that a significant sub-population of models *do* represent their target systems and we must understand how representation works, how models secure reference, and how models are more of less accurate in said representations to have a complete philosophical understanding of models. But, insofar as models serve functions that cannot be reduced down to facts about representation, understanding how models represent will not be sufficient to understand models *simpliciter*.

(Having played my realist cards, I also will say that everything I've said here can be co-opted, with minor tweaks, by a staunch instrumentalist. This will need to involve a denial or a deflation of the primary explanatory function of models which could be a natural way to pursue my project along anti-realist lines. While I do not seek to argue this, I'm content with anyone who will accept what I've said here, regardless of their purposes. I'm happy to sell hope to both sides.)

Onto the new program. There are three primary questions to which I give preliminary answers in an effort to begin building the foundation for this new way to think about models.

1. What non-representational terms, concepts, and tools will be needed?
2. How should we re-conceive past problems in the light of the explanatory-exploratory codependency in models?
3. What new questions should we ask?

To the first question, understanding how models work in terms of prescribing experiments is perhaps the most important concept for looking at the non-representational functions that models serve. The codependency of secondary explanation and exploration is most evidently seen in the experimental use of models, as I've argued throughout the paper.

Starting from the point of experimentation, we can begin to see a litany of different tools and concepts that will help analyze these functions in models. This will include measurement, data analytics, manipulations, controls, procedures, replicability, random assignment and selection, and more. None of these concepts are associated with those traditionally used in the representationalist camp and should be promising tools in this new foray of the philosophy of models.

To the second question, I only focus on two particularly salient problems which may become more soluble by looking at how models explain and explore non-representationally. I do this mainly for space reasons. However, there are a range of other past and contemporary issues surrounding models which may deserve to be looked at in light of what I've argued here. The two I have in mind are the *Problem of Inconsistent Models* and what I would call the *Approximation/Idealization/Abstraction Problem for Models*:

> *Problem of Inconsistent Models*: If different models purport to veridically represent relevant properties of the same target system and these properties are both essential and inconsistent

> with one another, then how can they both be successful? How can they both be models of the same system?
>
> *Approximation/Idealization/Abstraction Problem for Models*: How do models use approximation, idealization, and abstraction to function? How do these features help contribute to the success of models?

While I do not propose full answers to either issues here, I suggest some ways that headway can be made on them in this new research program. To the former, it's already been noted that the problem only remains when we try to understand models as completely representing their target systems (Massimi 2018). Traditionally, the response to this issue is to take a sort of holistic instrumentalism about models, which is unpalatable to many. Instead, we may be able to understand the sorts of success that these models have, particularly in the articulation and development of theory, in terms of secondary explanation and exploration. As I mentioned at the beginning of this subsection, we need not take an overarching instrumentalist approach about models to do so. And we may be able to better understand how parts of such models (or their successors) successfully represent their target systems through past experimentation.

In response to the *Approximation/Idealization/Abstraction Problem for Models*, it's widely agreed upon that these three qualities are central to the understanding of models. Of course, since these are explicit distortions of the target system made by a model, we run into immediate trouble if we try to understand them in the first instance as representing the target system. Quite obviously, approximations, idealizations, and abstractions in models *cannot* represent their target systems. These distortions, however, are purposeful. They are often instrumental in a variety of the areas in which we evaluate the success of a model. Much work can and should be done in seeing how these three techniques are operative in secondary explanations and exploration, particularly when helping scientists overcome practical obstacles in experimentation.

Finally, to the third question, I leave rather open-ended aside from a few remarks. For one, a deeper and more complex account of what scientific explorations are and how they are accomplished by models is needed and would be buoyed by the codependency I've drawn out here.

While what I have presented about them here should be sufficient for establishing their importance in models, much more deserves to be answered. Are there different ways that different types of models explain and explore? Is there any sort of trade-off between these two functions? Are there any structures in models that do not partake in these two functions?

Building off this, I've also gestured to the distinction between primary explanation and secondary explanation not being a categorical, but one of degree. Pursuing this thread is at once interesting for both philosophy of explanation and models. An even more provocative suggestion, given their codependency, is whether there is any sharp distinction between exploration and secondary explanation. Further research into these aspects of scientific inquiry, most of which would require philosophizing without representation, is rather enticing. From these two thoughts, we might even draw questions about the validity of the tripartite distinction between the introduction, colligation, and justification of theory going back to our discussion of scientific discovery, Whewell, and Reichenbach. Given how these functions seem to map onto this distinction, and given that they may differ only by degree, what does this say about the distinction? Only a development and pursuit of this research program will tell.

My final thesis is thus revisionary and prescriptive. As outlined at the beginning of the paper, much focus of late has been on how models represent. I think it is fair to say that many philosophers of science conceive of this as *the* problem of scientific models. I've argued this approach is misguided. Two major, codependent functions of significant class of models are to help us explain and explore their target systems. These functions are separable from a model's ability to represent, and thus cannot be understood in terms of representation. If we're to understand models, then, we should focus our attention on different problems pertaining to models, and answer these without appealing to representation. By revising our understanding of models and embracing this new thinking about them, I believe much more headway will be made on a philosophical theory of models.

References

Sylvain Bromberger. Why-questions. In Robert G. Colodny, editor, *Mind and Cosmos: Essays in Contemporary Science and Philosophy*, page 86–111. Pittsburgh, PA: University of Pittsburgh Press, 1966.

Otávio Bueno. Empirical adequacy: A partial structure approach. *Studies in the History and Philosophy of Science*, 28:585–610, 1997.

Otávio Bueno. Models and scientific representations. In P. Magnus and J. Busch, editors, *New Waves in Philosophy of Science*, page 94–111. Basingstoke: Palgrave Macmillan, 2010.

Otávio Bueno and Steven French. How theories represent. *The British Journal for the Philosophy of Science*, 62:857–94, 2011.

Craig Callender and Jonathan Cohen. There is no special problem about scientific representation. *Theoria*, 21:67–84, 2006.

Nancy Cartwright. *How the Laws of Physics Lie*. Oxford: Oxford University Press, 1983.

Mehmet Elgin and Elliott Sober. Cartwright on explanation and idealization. *Erkenntnis*, 57: 441–50, 2002.

Michael Faraday. Experimental researches in electricity. third series. *Philosophical Transactions of the Royal Society of London*, 123:23–54, 1833.

Grant Fisher. The autonomy of models and explanation: Anomalous molecular rearrangements in early twentieth-century physical organic chemistry. *Studies in History and Philosophy of Science Part A*, 37:562–84, 2006.

Steven French. A model-theoretic account of representation (or, i don't know much about art. . . but i know it involves isomorphism). *Philosophy of Science*, 70:1472–83, 2003.

Roman Frigg. Fiction and scientific representation. In Roman Frigg and Matthew C. Hunter, editors, *Beyond Mimesis and Convention: Representation in Art and Science*, page 97–138. Berlin and New York: Springer, 2010a.

Roman Frigg. Models and fiction. *Synthese*, 172:251–68, 2010b.

Roman Frigg and James Nguyen. The fiction view of models reloaded. *The Monist*, 99:225–42, 2016.

Roman Frigg and James Nguyen. *Modelling Nature. An Opinionated Introduction to Scientific Representation*. New York: Springer, 2020.

Roman Frigg and James Nguyen. Seven myths about the fiction view of models. In Alejandro Casini and Juan Redmond, editors, *Models and Idealizations in Science: Artifactual and Fictional Approaches*, page 133–157. Cham: Springer, 2021.

Axel Gelfert. *How to Do Science with Models: A Philosophical Primer*. Cham: Springer International Publishing, 2016.

Ronald N. Giere. *Explaining Science: A Cognitive Approach*. Chicago: Chicago University Press, 1988.

Ronald N. Giere. How models are used to represent reality. *Philosophy of Science*, 71:742–52, 2004.

Peter Godfrey-Smith. The strategy of model-based science. *Biology and Philosophy*, 21:725–40, 2006.

Norwood Russell Hanson. Is there a logic of scientific discovery? *Australasian Journal of Philosophy*, 38:91–106, 1960.

Carl G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press, 1965.

Thomas Kuhn. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 2 edition, 1970.

Larry Laudan. Why was the logic of discovery abandoned? In Thomas Nickles, editor, *Scientific Discovery*, volume I, pages 173–83. Dordrecht: D. Reidel, 1980.

Lorenzo Magnani. Creative abduction and hypothesis withdrawal. In Joke Meheus and Thomas Nickles, editors, *Models of Discovery and Creativity*. Dordrecht: Springer, 2009.

Michela Massimi. Perspectival modeling. *Philosophy of Science*, 85:335–59, 2018.

Michela Massimi. Two kinds of exploratory models. *Philosophy of Science*, 86:869–81, 2019.

Thomas Nickles. Positive science and discoverability. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, page 13–27, 1984.

Thomas Nickles. Beyond divorce: Current status of the discovery debate. *Philosophy of Science*, 52:177–206, 1985.

Thomas Nickles. Truth or consequences? generative versus consequential justification in science. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, page 393–405, 1989.

Steven French Otávio Bueno and James Ladyman. On representing the relationship between the mathematical and the empirical. *Philosophy of Science*, 69:452–73, 2002.

Karl Popper. *The Logic of Scientific Discovery*. London: Hutchinson, 1959.

Karl Popper. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Hutchinson, 1963.

Hans Reichenbach. *Experience and Prediction. An Analysis of the Foundations and the Structure of Knowledge*. Chicago: The University of Chicago Press, 1938.

Ernest Rutherford. The scattering of α and β particles by matter and the structure of the atom. *Philosophical Magazine Series 6*, 21:669–88, 1911.

Fiora Salis. The new fiction view of models. *The British Journal for the Philosophy of Science*, 72:717–42, 2021.

Kenneth Schaffner. *Discovery and Explanation in Biology and Medicine*. Chicago: The University of Chicago Press, 1993.

Michael Strevens. The causal and unification approaches to explanation unified—causally. *Noûs*, 38:154–76, 2004.

Michael Strevens. *Depth: An Account of Scientific Explanation*. Cambridge, MA, and London: Harvard University Press, 2008.

Mauricio Suárez. Scientific representation: Against similarity and isomorphism. *International Studies in the Philosophy of Science*, 17:225–44, 2003.

Mauricio Suárez. An inferential conception of scientific representation. *Philosophy of Science*, 71: 767–79, 2004.

Mauricio Suárez. Deflationary representation, inference, and practice. *Studies in History and Philosophy of Science*, 49:36–47, 2015.

Paul Thagard. Conceptual combination and scientific discovery. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pages 3–12, 1984.

Paul Thagard. *How Scientists Explain Disease*. Princeton: Princeton University Press, 1999.

Paul Thagard. How brains make mental models. In *Model-Based Reasoning in Science and Technology*, pages 447–61. Berlin and Heidelberg: Springer, 2010.

Paul Thagard. *The Cognitive Science of Science*. Cambridge, MA: MIT Press, 2012.

J.J. Thomson. On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. *Philosophical Magazine Series 6*, 7:237–65, 1904.

Bas van Fraassen. *The Scientific Image*. Oxford: Oxford University Press, 1980.

Bas van Fraassen. *Scientific Representation: Paradoxes of Perspective*. Oxford: Oxford University Press, 2008.

William Whewell. *The Philosophy of the Inductive Sciences*, volume II. London: Routledge/Thoemmes, 1840.

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