1. The Ontology of Models

Axel Gelfert

The term *scientific model* picks out a great many things, including scale models, physical models, sets of mathematical equations, theoretical models, toy models, and so forth. This raises the question of whether a general answer to the question *What is a model*? is even possible. This chapter surveys a number of philosophical approaches that bear on the question of what, in general, a scientific model is. While some approaches aim for a unitary account that would apply to models in general, regardless of their specific features, others take as their basic starting point the manifest heterogeneity of models in scientific practice. This chapter first motivates the ontological question of what models are by reflecting on the diversity of different kinds of models and arguing that models are best understood as *functional entities*. It then provides some historical background regarding the use of analogy in science as a precursor to contemporary notions of *scientific model*. This is followed by a contrast between the syntactic and the semantic views of theories and models and their different stances toward the question of what a model is. Scientists, too, typically operate with tacit assumptions about the ontological status of models: this gives rise to what has been called the *folk ontology* of models, according to which models may be thought of as descriptions of missing (i. e., uninstantiated) systems. There is a close affinity between this view and recent philosophical positions (to be discussed in the penultimate section) according to which models are fictions. This chapter concludes by considering various pragmatic conceptions of models, which are typically associated with what may be called *mixed ontologies*, that is, with the view that any quest for a unitary account of the nature of models is bound to be fruitless.

1.1 **Kinds of Models: Examples from Scientific Practice** 6
1.2 **The Nature and Function of Models** 8
1.3 **Models as Analogies and Metaphors** 10
1.4 **Models Versus the Received View: Sentences and Structures** 12
1.4.1 Models and the Study of Formal Languages 13
1.4.2 The Syntactic View of Theories 13
1.4.3 The Semantic View 14
1.4.4 Partial Structures 15
1.5 **The Folk Ontology of Models** 16
1.6 **Models and Fiction** 18
1.7 **Mixed Ontologies: Models as Mediators and Epistemic Artifacts** 20
1.7.1 Models as Mediators 20
1.7.2 Models as Epistemic Artifacts 21
1.8 **Summary** 21
References 22

The philosophical discussion about models has emerged from a cluster of concerns, which span a range of theoretical, formal, and practical questions across disciplines ranging from logic and mathematics to aesthetics and artistic representations. In what follows, the term *models* will normally be taken as synonymous to *scientific models*, and any departure from this usage – for example, when discussing the use of models in nonscientific settings – will either be indicated explicitly or will be clear from context. Focusing on scientific models helps to clarify matters, but still leaves a wide range of competing philosophical approaches for discussion. This chapter will summarize and critically discuss a number of such approaches, especially those that shed light on the question *what is a model*?; these will range from views that, by now, are of largely historical interest to recent proposals at the cutting edge of the philosophy of science. While the emphasis throughout will be on the ontology of models, it will often be necessary to also reflect on their function, use, and construction. This is not meant to duplicate the discussion provided in other chapters of this handbook; rather, it is the natural result of scientific models having traditionally been defined either in terms of their function (e.g., to provide representations of target systems) or via their relation to other (purportedly) better understood entities, such as scientific theories.

The rest of this chapter is organized as follows: Sect. 1.1 will set the scene by introducing a number of examples of scientific models, thereby raising the question of what degree of unity any philosophical account of scientific models can reasonably aspire to. Section 1.2 will characterize models as functional entities and will provide a general taxonomy for how to classify various possible philosophical approaches. A first important class of specific accounts, going back to nineteenth-century scientists and philosophers, will be discussed in Sect. 1.3, which focuses on models as analogies. Section 1.4 is devoted to formal approaches that dominated much of twentieth-century discussion of scientific models. In particular, it will survey the syntactic view of theories and models and its main competitor, the semantic view, along with recent formal approaches (such as the partial structures approach) which aim to address the shortcomings of their predecessors. Section 1.5 provides a sketch of what has been called the *folk ontology* of models – that is, a commonly shared set of assumptions that inform the views of scientific practitioners. On this view, models are place-holders for *imaginary concrete systems* and as such are not unlike fictions. The implications of fictionalism about models are discussed in Sect. 1.6. Finally, in Sect. 1.7, recent pragmatic accounts are discussed, which give rise to what may be called a *mixed ontology*, according to which models are best conceived of as a heterogeneous mixture of elements.

1.1 Kinds of Models: Examples from Scientific Practice

Models can be found across a wide range of scientific contexts and disciplines. Examples include the Bohr model of the atom (still used today in the context of science education), the billiard ball model of gases, the DNA double helix model, scale models in engineering, the Lotka–Volterra model of predator–prey dynamics in population biology, agent-based models in economics, the Mississippi River Basin model (which is a 200 acres hydraulic model of the waterways in the entire Mississippi River Basin), and general circulation models (GCMs), which allow scientists to run simulations of Earth's climate system. The list could be continued indefinitely, with the number of models across the natural and social sciences growing day by day.

In philosophical discussions of scientific models, the situation is hardly any different. The *Stanford Encyclopedia of Philosophy* gives the following list of model types that have been discussed by philosophers of science [1.1]:

> "Probing models, phenomenological models, computational models, developmental models, explanatory models, impoverished models, testing models, idealized models, theoretical models, scale models, heuristic models, caricature models, didactic models, fantasy models, toy models, imaginary models, mathematical models, substitute models, iconic models, formal models, analogue models and instrumental models."

The proliferation of models and model types, in the sciences as well as in the philosophical literature, led *Goodman* to lament in his 1968 *Languages of Art* [1.2, p. 171]: "Few terms are used in popular and scientific discourse more promiscuously than *model*." If this was true of science and popular discourse in the late 1960s, it is all the more true of the twenty-first century philosophy of science.

As an example of a physics-based model, consider the *Ising model*, proposed in 1925 by the German physicist Ernst Ising as a model of ferromagnetism in certain metals. The model starts from the idea that a macroscopic magnet can be thought of as a collection of elementary magnets, whose orientation determines the overall magnetization. If all the elementary magnets are aligned along the same axis, then the system will be perfectly ordered and will display a maximum value of the magnetization. In the simplest one-dimensional (1-D) case, such a state can be visualized as a chain of *elementary magnets*, all pointing the same way

$$\cdots\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\cdots$$

The alignment of elementary magnets can be brought about either by a sufficiently strong external magnetic field or it can occur spontaneously, as will happen below a critical temperature, when certain substances (such as iron and nickel) undergo a ferromagnetic phase transition. Whether or not a system will undergo a phase transition, according to thermodynamics, depends on its energy function, which in turn is determined by the interactions between the component parts of the system. For example, if neighboring *elementary magnets* interact in such a way as to favor alignment, there is a good chance that a spontaneous phase transition may occur below a certain temperature. The energy function, then, is crucial to the model and, in the case of the Ising model, is defined as

$$E = -\sum_{i,j} J_{ij} S_i S_j \,,$$

with the variable S_i representing the orientation ($+1$ or -1) of an elementary magnet at site i in the crystal lattice and J_{ij} representing the strength of interaction between two such elementary magnets at different lattice sites i and j.

Contrast this with *model organisms* in biology, the most famous example of which is the fruit fly *Drosophila melanogaster*. Model organisms are real organisms – actual plants and animals that are alive and can reproduce – yet they are used as representations either of another organism (e.g., when rats are used in place of humans in medical research) or of a biological phenomenon that is more universal (e.g., when fruit flies are used to study the effects of crossover between homologous chromosomes). Model organisms are often bred for specific purposes and are subject to artificial selection pressures, so as to purify and *standardize* certain features (e.g., genetic defects or variants) that would not normally occur, or would occur only occasionally, in populations in the wild. As *Ankeny* and *Leonelli* put it, in their ideal form "model organisms are thought to be a relatively simplified form of the class of organism of interest" [1.3, p. 318]; yet it often takes considerable effort to work out the actual relationships between the model organism and its target system (whether it be a certain biological phenomenon or a specific class of target organisms). Tractability and various experimental desiderata – for example, a short life cycle (to allow for quick breeding) and a relatively small and compact genome (to allow for the quick identification of variants) – take precedence over theoretical questions in the choice of model organisms; unlike for the Ising model, there is no simple mathematical formula that one can rely on to study how one's model behaves, only the messy world of real, living systems.

The Ising model of ferromagnetism and model organisms such as *Drosophila melanogaster* may be at opposite ends of the spectrum of scientific models. Yet the diversity among those models that occupy the middle ground between theoretical description and experimental system is no less bewildering. How, one might wonder, can a philosophical account of scientific models aspire to any degree of unity or generality in the light of such variety? One obvious strategy is to begin by drawing distinctions between different overarching types of models. Thus, *Black* [1.4] distinguishes between four such types:

1. Scale models
2. Analog models
3. Mathematical models
4. Theoretical models.

The basic idea of scale and analog models is straightforward: a scale model increases or decreases certain (e.g., spatial) features of the target system, so as to render them more manageable in the model; an analog model also involves the change of medium (as in once popular hydraulic models of the economy, where the flow of money was represented by the flow of liquids through a system of pumps and valves). Mathematical models are constructed by first identifying a number of relevant variables and then developing empirical hypotheses concerning the relations that may hold between the variables; through (often drastic) simplification, a set of mathematical equations is derived, which may then be evaluated analytically or numerically and tested against novel observations. Theoretical models, finally, begin usually by extrapolating imaginatively from a set of observed facts and regularities, positing new entities and mechanisms, which may be integrated into a possible theoretical account of a phenomenon; comparison with empirical data usually comes only at a later stage, once the model has been formulated in a coherent way.

Achinstein [1.5] includes mathematical models in his definition of *theoretical model*, and proposes an analysis in terms of sets of assumptions about a model's target system. This allows him to include Bohr's model of the atom, the DNA double-helix model (considered as a set of structural hypotheses rather than as a physical ball-and-stick model), the Ising model, and the Lotka–Volterra model among the class of theoretical systems. Typically, when a scientist constructs a theoretical model, she will help herself to certain established principles of a more fundamental theory to which she is committed. These will then be adapted or modified, notably by introducing various new assumptions specific to the case at hand. Typically, an inner structure or mechanism is posited which is thought to explain the features of the target system. At the same time, there is the (often explicit) acknowledgment that the target system is far more complex than the model is able to capture: in this sense, a theoretical model is believed by the practitioner to be false as a description of the target system. However, this acknowledgment of the limits of applicability of models also allows researchers to simultaneously use different models of the same target system alongside each other. Thus understood, theoretical models usually involve the combination of general theoretical principles and specific auxiliary assumptions, which may only be valid for a narrow range of parameters.

1.2 The Nature and Function of Models

The great variety of models employed in scientific practice, as illustrated by the long list given in the preceding section, suggests two things. First, it makes vivid just how central the use of models is to the scientific enterprise and to the self-image of scientists. As *von Neumann* put it, with some hyperbole [1.6, p. 492]: "The sciences do not try to explain, they hardly even try to interpret, they mainly make models." Whatever shape and form the scientific enterprise might take without the use of models, it seems safe to say that it would not look anything like science as we presently know it. Second, one might wonder whether it is at all reasonable to look for a unitary philosophical account of models. Given the range of things we call *models*, and the diversity of uses to which they are being put, it may simply not be possible to give a one-size-fits-all answer to the question *what is a model?* This has led some commentators to propose quietism as the only viable attitude toward ontological questions concerning models and theories. As *French* puts it [1.7, p. 245],

> "whereas positing the reality of quarks or genes may contribute to the explanation of certain features of the physical world, adopting a similar approach toward theories and models – that is, reifying them as entities for which a single unificatory account can be given – does nothing to explain the features of scientific practice."

While there are good grounds for thinking that quietism should only be a position of last resort in philosophy, the sentiment expressed by French may go some way toward explaining why there has been a relative dearth of philosophical work concerning the ontology of models. The neglect of ontological questions concerning models has been remarked upon by a number of contributors, many of whom, like *Contessa*, find it [1.8, p. 194]

> "surprising if one considers the amount of interest raised by analogous questions about the ontology and epistemology of mathematical objects in the philosophy of mathematics."

A partial explanation of this discrepancy lies in the arguably greater heterogeneity in what the term *scientific models* is commonly thought to refer to, namely, anything from physical ball-and-stick models of chemical molecules to mathematical models formulated in terms of differential equations. (If we routinely included dividers, compasses, set squares, and other technical drawing tools among, say, the class of *geometrical entities*, the ontology of mathematical entities, too, would quickly become rather unwieldy!)

In the absence of any widely accepted unified account of models – let alone one that would provide a conclusive answer to ontological questions arising from models – it may be natural to assume, as indeed many contributors to the debate have done, that "if all scientific models have something in common, this is not their *nature* but their *function*" [1.8, p. 194]. One option would be to follow the quietist strategy concerning the ontology of models and "refuse to engage with this issue and ask, instead, how can we best represent these features [and functions of models] in order that we can understand" [1.7, p. 245] the practice of scientific modeling. Alternatively, however, one might simply accept that the function of models in scientific inquiry is our best – and perhaps only – guide when exploring answers to the question *what is a model?*. At the very least, it is not obvious that an exploration of the ontological aspects of models is necessarily fruitless or misguided. *Ducheyne* puts this nicely when he argues that [1.9, p. 120],

> "if we accept that models are functional entities, it should come as no surprise that when we deal with scientific models ontologically, we cannot remain silent on how such models function as carriers of scientific knowledge."

As a working assumption, then, let us treat scientific models as *functional entities* and explore how much ontological unity – over and above their *mere* functional role – we can give to the notion of *scientific model*.

Two broad classes of functional characterizations of models can be distinguished, according to which it is either *instantiation* or *representation* that lie at the heart of how models function. As *Giere* [1.10] sees it, on the *instantial view*, models instantiate the axioms of a theory, where the latter is understood as being comprised of linguistic statements, including mathematical statements and equations. (For an elaboration of how such an account might turn out, see Sect. 1.4.) By contrast, on the *representational view*, "language connects not directly with the world, but rather with a model, whose characteristics may be precisely defined"; the model then connects with the world "by way of similarity between a model and the designated parts of the world" [1.10, p. 156]. Other proponents of the representational view have de-emphasized the role of similarity, while still endorsing representation as one of the key functions of scientific models. Generally speaking, proponents of the representational view consider models to be "tools for *representing the world*," whereas those who favor the instantial view regard them primarily as "providing a means for interpreting formal systems" [1.10, p. 44].

Within the class of representational views, one can further distinguish between views that emphasize the *informational* aspects of models and those that take their *pragmatic* aspects to be more central. *Chakravartty* nicely characterizes the informational variety of the representational view as follows [1.11, p. 198]:

> "The idea here is that a scientific representation is something that bears an objective relation to the thing it represents, on the basis of which it contains information regarding that aspect of the world."

The term *objective* here simply means that the requisite relation obtains independently of the model user's beliefs or intentions as well as independently of the specific representational conventions he or she might be employing. Giere's similarity-based view of representation – according to which scientific models represent in virtue of their being similar to their target systems in certain specifiable ways – would be an example of such an informational view similarity, as construed by Giere, is a relation that holds between the model and its target, irrespective of a model user's beliefs or intentions, and regardless of the cognitive uses to which he or she might put the model. Other philosophical positions that are closely aligned with the informational approach might posit that, for a model to represent its target, the two must stand in a relation of isomorphism, partial isomorphism, or homomorphism to one another.

By contrast, the *pragmatic* variety of the representational view of models posits that models function as representations of their targets in virtue of the cognitive uses to which human reasoners put them. The basic idea is that a scientific model facilitates certain cognitive activities – such as the drawing of inferences about a target system, the derivation of predictions, or perhaps a deepening of the scientific understanding – on the part of its user and, therefore, necessarily involves the latter's cognitive interests, beliefs, or intentions. *Hughes* [1.12], for example, emphasizes the interplay of three cognitive–theoretical processes – denotation, demonstration, and interpretation – which jointly give rise to the representational capacity of (theoretical) models in science. On Hughes' (aptly named) *DDI account* of model-based representation, *denotation* accounts for the fact that theoretical elements of a model purport to refer to elements in the physical world. The possibility of *demonstration* from within a model – in particular, the successful mathematical derivation of results for models that lend themselves to mathematical derivation techniques – attests both to the models having a nontrivial internal dynamic and to its being a viable object of fruitful theoretical investigation. Through successful *interpretation*, a model user then relates the theoretically derived results back to the physical world, including the model's target system. Clearly, the DDI account depends crucially on there being someone who engages in the activities of interpreting and demonstrating – that is, it depends on the cognitive activities of human agents, who will inevitably draw on their background knowledge, cognitive interests, and derivational skills in establishing the requisite relations for bringing about representation.

The contrast between informational and pragmatic approaches to model-based representation roughly maps onto another contrast, between what *Knuuttila* has dubbed *dyadic* and *triadic* approaches. The former takes "the model–target dyad as a basic unit of analysis concerning models and their epistemic values" [1.13, p. 142]. This coheres well with the informational approach which, as discussed, tends to regard models as (often abstract) structures that stand in a relation of isomorphism, or partial isomorphism, to a target system. By contrast, triadic accounts – in line with pragmatic views of model-based representation – based representation shift attention away from models and the abstract relations they stand in, toward modeling as a theoretical activity pursued by human agents with cognitive interests, intentions, and beliefs. On this account, model-based representation cannot simply be a matter of any abstract relationship between the model and a target system since one cannot, as *Suárez* puts it, "reduce the essentially intentional judgments of representation users to facts about the source and target object or systems and their properties" [1.14, p. 768]. Therefore, so the suggestion goes, the model–target dyad needs to be replaced by a three-place relation between the model, its target, and the model user. Suárez, for example, proposes an inferentialist account of model-based representation, according to which a successful model must allow "competent and informed agents to draw specific inferences regarding" [1.14, p. 773] the target system – thereby making the representational success of a model dependent on the qualities of a (putative) model user.

1.3 Models as Analogies and Metaphors

Some scholars trace the emergence of the concept of a *scientific model* to the second half of the nineteenth century [1.15]. Applying our contemporary concept of *model* to past episodes in the history of science, we can of course identify prior instances of models being employed in science; however, until the nineteenth century scientists were engaged in little systematic self-reflection on the uses and limitations of models. Philosophy of science took even longer to pay attention to models in science, focusing instead on the role and significance of scientific theories. Only from the middle of the twentieth century onward did philosophical interest in models acquire the requisite momentum to carry the debate forward. Yet in both science and philosophy, the term *model* underwent important transformations, so it will be important to identify some of these shifts, in order to avoid unnecessary ambiguity and confusion in our exploration of the question *What is a model?*.

Take, for example, *Duhem*'s dismissal, in 1914, of what he takes to be the excessive use of models in Maxwell's theory of electromagnetism, as presented in an English textbook published at the end of the nineteenth century [1.16, p. 7]:

> "Here is a book intended to expound the modern theories of electricity and to expound a new theory. In it there are nothing but strings which move round pulleys which roll around drums, which go through pearl beads, which carry weights; and tubes which pump water while others swell and contract; toothed wheels which are geared to one another and engage hooks. We thought we were entering the tranquil and neatly ordered abode of reason, but we find ourselves in a factory."

What Duhem is mocking in this passage, which is taken from a chapter titled *Abstract Theories and Mechanical Models*, is a style of reasoning that is dominated by the desire to *visualize* physical processes in purely mechanical terms. His hostility is thus directed at *mechanical* models only – as the implied contrast in the chapter title makes clear – and does not extend to the more liberal understanding of the term *scientific model* in philosophy of science today.

Indeed, when it comes to the use of *analogy* in science, Duhem is much more forgiving. The term *analogy*, which derives from the Greek expression for *proportion*, itself has multiple uses, depending on whether one considers its use as a rhetorical device or as a tool for scientific understanding. Its general form is that of "pointing to a resemblance between relations in two different domains, that is, A is related to B like C is related to D" [1.17, p. 110]. An analogy may be considered merely *formal*, when only the relations (but not the relata) resemble another, or it may be *material*, when the relata from the two domains (i. e., A and B on one side, C and D on the other) have certain attributes or characteristics in common. *Duhem*'s understanding of *analogy* is more specific, in that he conceives of analogy as being a relation between two sets of statements, such as between one theory and another [1.16, p. 97]:

> "Analogies consist in bringing together two abstract systems; either one of them already known serves to help us guess the form of the other not yet known, or both being formulated, they clarify the other. There is nothing here that can astonish the most rigorous logician, but there is nothing either that recalls the procedures dear to ample but shallow minds."

Consider the following example: When Christiaan Huygens (1629–1695) proposed his theory of light, he did so on the basis of *analogy* with the theory of sound waves: the relations between the various attributes and characteristics of light are similar to those described by acoustic theory for the rather different domain of sound. Thus understood, analogy becomes a legitimate instrument for learning about one domain on the basis of what we know about another. In modern parlance, we might want to say that sound waves provided Huygens with a good *theoretical model* – at least given what was known at the time – for the behavior of light.

There is, however, a risk of ambiguity in that last sentence – an ambiguity which, as *Mellor* [1.18, p. 283] has argued, it would be wrong to consider harmless. Saying that *sound waves provide a good model for the theory of light* appears to equate the model *with the sound waves* – as though one physical object (sound waves) could be identified with the model. At first sight, this might seem unproblematic, given that, as far as wave-like behavior is concerned, we do take light and sound to be relevantly analogous. However, while it is indeed the case that "some of the constructs called *analogy* in the nineteenth century would today be routinely referred to as *models*" [1.19, p. 46], it is important to distinguish between, on the one hand, *analogy* as the similarity relation that exists between a theory and another set of statements and, on the other hand, the latter set of statements as the *analog of the theory*. Furthermore, we need to distinguish between the analog (e.g., the theory of sound waves, in Huygens's case) and the set of entities *of which the analog is true* (e.g., the sound waves themselves). (On this point, see [1.18, p. 283].) What Duhem resents about the naïve use of what he refers to as *mechanical models* is the hasty conflation of the visualized entities – (imaginary) pulleys, drums, pearl beads, and toothed wheels – with what is *in fact* scientifically valuable, namely the relation of analogy that exists between, say, the theory of light and the theory of sound.

This interpretation resolves an often mentioned tension – partly perpetuated by Duhem himself, through his identification of different styles of reasoning (the *English* style of physics with its emphasis on mechanical models, and the *Continental* style which prizes mathematical principles above all) – between Duhem's account of models and that of the English physicist Norman Campbell. Thus, *Hesse*, in her seminal essay *Models and Analogies in Science* [1.20], imagines a dialogue between a *Campbellian* and a *Duhemist*. At the start of the dialogue, the Campbellian attributes to the Duhemist the following view: "I imagine that along with most contemporary philosophers of science, you would wish to say that the use of models or analogs is not essential to scientific theorizing and that [. . .] the theory as a whole does not require to be interpreted by means of any model." To this, the Duhemist, who admits that "models may be useful guides in suggesting theories," replies: "When we have found an acceptable theory, any model that may have led us to it can be thrown away. Kekulé is said to have arrived at the structure of the benzene ring after dreaming of a snake with its tail in its mouth, but no account of the snake appears in the textbooks of organic chemistry." The Campbellian's rejoinder is as follows: "I, on the other hand, want to argue that models in some sense are essential to the logic of scientific theories" [1.20, pp. 8–9]. The quoted part of Hesse's dialogue has often been interpreted as suggesting that the bone of contention between Duhem and Campbell is the status of *models in general* (in the modern sense that includes theoretical models), with Campbell arguing in favor and Duhem arguing against. But we have already seen that Duhem, using the language of *analogy*, *does* allow for theoretical models to play an important role in science. This apparent tension can be resolved by being more precise about the target of Duhem's criticism: "Kekulé's snake dream might illustrate the use of a visualizable model, but it certainly does not illustrate the use of an analogy, in Duhem and Campbell's sense" [1.18, p. 285]. In other words, Duhem is not opposed to scientific models in general, but to its mechanical variety in particular. And, on the point of over-reliance on mechanical models, *Campbell*, too, recognizes that dogmatic attachment to such a style of reasoning is *open to criticism*. Such a dogmatic view would hold "that theories are completely satisfactory only if the analogy on which they are based is mechanical, that is to say, if the analogy is with the laws of mechanics" [1.21, p. 154]. Campbell is clearly more sympathetic than Duhem toward our "craving for mechanical theories," which he takes to be firmly rooted in our psychology. But he insists that [1.21, p. 156]

> "we should notice that the considerations which have been offered justify only the attempt to adopt some form of theory involving ideas closely related to those of force and motion; it does not justify the attempt to force all such theories into the Newtonian mold."

To be sure, significant differences between Duhem and Campbell remain, notably concerning what *kinds* of uses of analogies in science (or, in today's terminology, of scientific – including theoretical – models) are appropriate. For Duhem, such uses are limited to a heuristic role in the discovery of scientific theories. By contrast, *Campbell* claims that "in order that a theory may be valuable [. . .] it must display analogy" [1.21, p. 129] – though it should be emphasized again, not necessarily analogy *of the mechanical sort*. (As *Mellor* argues, Duhem and Campbell differ chiefly in their views of scientific theories and less so in their take on analogy, with Duhem adopting a more *static* perspective regarding theories and Campbell taking a more realist perspective [1.18].)

It should be said, though, that Hesse's *Campbellian* and *Duhemist* are at least partly intended as caricatures and serve as a foil for Hesse's own account of models as analogies. The account hinges on a three-part distinction between *positive*, *negative*, and *neutral* analogies [1.20]. Using the billiard ball model of gases as her primary example, Hesse notes that some characteristics are shared between the billiard balls and the gas atoms (or, rather, are ascribed by the billiard ball model to the gas atoms); these include velocity, momentum, and collision. Together, these constitute the *positive* analogy. Those properties we know to belong to billiard balls, but not to gas atoms – such as color – constitute the *negative* analogy of the model. However, there will typically be properties of the model (i. e., the billiard ball system) of which we do not (yet) know whether they also apply to its target (in this case, the gas atoms). These form the *neutral* analogy of the model. Far from being unimportant, the neutral analogy is crucial to the fruitful use of models in scientific inquiry, since it holds out the promise of acquiring new knowledge about the target system by studying the model in its place [1.20, p. 10]:

> "If gases are really like collections of billiard balls, except in regard to the known negative analogy, then from our knowledge of the mechanics of billiard balls, we may be able to make new predictions about the expected behavior of gases."

In dealing with scientific models we may choose to disregard the negative analogy (which results in what Hesse calls *model*$_1$) and consider only the known positive and neutral analogies – that is, only those properties that are shared, or for all we know may turn out to be shared, between the target system and its analog. (On the terminology discussed in Sect. 1.1, due to Black and Achinstein, model$_1$ would qualify as a *theoretical model*.) This, Hesse argues, typically describes our use of models for the purpose of explanation: we resolve to treat model$_1$ as taking the place of the phenomena themselves. Alternatively, we may actively include the negative analogy in our considerations, resulting in what Hesse calls *model*$_2$ or a form of analog model. Given that, let us assume, the model system (e.g., the billiard balls) was chosen because it was observable – or, at any rate, more accessible than the target system (e.g., the gas) – model$_2$ allows us to study the similarities and dissimilarities between the two analogous domains; model$_2$, qua being a model for its target, thus has a deeper structure than the system of billiard balls considered in isolation – and, like model$_1$, importantly includes the neutral analogy, which holds out the promise of novel insights and predictions. As *Hesse* puts it, in the voice of her Campbellian interlocutor [1.20, pp. 12–13]:

> "My whole argument is going to depend on these features [of the neutral analogy] and so I want to make it clear that I am not dealing with static and formalized theories, corresponding only to the known positive analogy, but with theories in the process of growth."

Models have been discussed not only in terms of analogy, but also in terms of metaphor. *Metaphor*, more explicitly than *analogy*, refers to the linguistic realm: a metaphor is a linguistic expression that involves at least one part that is being transferred from a domain of discourse where it is common to another – the target domain – where it is uncommon. The existence of an analogy may facilitate such a transfer of linguistic expression; at the same time, it is entirely possible that "it is the metaphor that prompts the recognition of analogy" [1.17, p. 114] – both are compatible with one another and neither is obviously prior to the other. Metaphorical language is widespread in science, not just in connection with models: for example, physicists routinely speak of *black holes* and *quantum tunneling* as important predictions of general relativity theory and quantum theory, respectively. Yet, as *Soskice* and *Harré* note, there is a special affinity between models and metaphor [1.22, p. 302]:

> "The relationship of model and metaphor is this: if we use the image of a fluid to explicate the supposed action of the electrical energy, we say that the fluid is functioning as a model for our conception of the nature of electricity. If, however, we then go on to speak of the *rate of flow* of an *electrical current*, we are using metaphorical language based on the fluid model."

In spite of this affinity, it would not be fruitful to simply equate the two – let alone jump to the conclusion that, in the notion of *metaphor*, we have found an answer to the question *What is a model?*. Models and metaphors both issue in descriptions, and as such they may draw on analogies we have identified between two otherwise distinct domains; more, however, needs to be said about the nature of the relations that need to be in place for something to be considered a (successful) model of its target system or phenomenon.

1.4 Models Versus the Received View: Sentences and Structures

Much of the philosophical debate about models is indebted to model theory as a branch of (first-order) mathematical logic. Two philosophical frameworks for thinking about scientific models and theories – the *syntactic view* of models and theories and its main competitor, the *semantic view* – can be traced back to these origins; they are the topic of this section. (For a more extensive discussion, see also other chapters in this handbook.) The syntactic view (Sect. 1.4.2) is closely aligned with logical positivism, which dominated much anglophone philosophy of science until the mid-1960s, and is sometimes referred to as *the received view*. Given that less rigid approaches and an overarching movement toward pluralism have reshaped the philosophy of science over the past half-century or so, this expression is somewhat dated; to make matters worse, other contributors to the debate have, over time, come to apply the same label to the syntactic view's main competitor, the semantic view of models and theories. Instead of adjudicating which position deserves this dubious honor, the present section will discuss how each view conceives of models. Before doing so, however, a few preliminaries are in order concerning the competing views' joint origins in logical model theory.

1.4.1 Models and the Study of Formal Languages

Model theory originated as the study of formal languages and their interpretations, starting from a Tarski-style truth theory based only on notions from syntax and set theory. On a broader understanding, the restriction to formal languages may be dropped, so as to include scientific languages (which are often closer to natural language than to logic), or even natural languages. However, the distinction between the syntax and the semantics of a language, which is sharpest in logic, also provides a useful framework for studying scientific languages and has guided the development of both the syntactic and the semantic views of theories and models. The *syntax* of a language L is made up of the vocabulary of L, along with the rules that determine which sequence of symbols counts as a well-formed expression in L; in turn, the *semantics* of L provides interpretations of the symbolic expressions in L, by mapping them onto another relational structure R, such that all well-formed expressions in L are rendered intelligible (e.g., via rules of composition) and can be assessed in terms of their truth or falsity in R.

The contrast between the syntax and the semantics of a language allows for two different approaches to the notion of a *theory*. A theory T may either be defined syntactically, as the set of all those sentences that can be derived, through a proper application of the syntactic rules, from a set of axioms (i. e., statements that are taken to be fundamental); or it may be defined semantically, as all those (first-order) sentences that a particular structure, M, satisfies. An example of the former would be Euclidean geometry, which consists of five axioms and all the theorems derivable from them using geometrical rules; an example of the latter would be group theory, which simply consists of all those first-order sentences that a set of groups – definable in terms of set-theoretic entities – satisfies. (This example, and much of the short summary in this section, is owed to [1.23]; for further discussion, see references therein.) The syntactic and semantic definitions of what a theory is are closely related: starting from the semantic definition, to see whether a particular structure M is a model of an axiomatizable first-order theory T, all that one needs to show is that M satisfies the axioms.

1.4.2 The Syntactic View of Theories

The syntactic view of theories originated from the combination of the insights – or, to put it a little more cautiously, fundamental tenets – of two research programs: the philosophical program, aligned with Pierre Duhem (Sect. 1.3) and Henri Poincaré, of treating (physical) theories as systems of hypotheses designed to *save the phenomena*, and the mathematical program, pioneered by David Hilbert, which sought to formalize (mathematical) theories as axiomatic systems. By combining the two, it seemed possible to identify a theory with the set of logical consequences that could be derived from its fundamental principles (which were to be treated as axioms), using only the rules of the language in which the theory was formulated. In spite of its emphasis on syntax, the syntactic view is not entirely divorced from questions of semantics. When it comes to scientific theories, we are almost always dealing with *interpreted* sets of sentences, some of which – the fundamental principles or axioms – are more basic than others, with the rest derivable using syntactic rules. The question then arises at which level interpretation of the various elements of a theory is to take place. This is where the slogan *to save the phenomena* points us in the right direction: on the syntactic view, interpretation only properly enters at the level of matching singular theoretical predictions, formulated in strictly observational terms, with the observable phenomena. Higher level interpretations – for example, pertaining to purely theoretical terms of a theory (such as posited unobservable entities, causal mechanisms, laws, etc.) – would be addressed through *correspondence rules*, which offered at least a partial interpretation, so that *some* of the meaning of such higher level terms of a theory could be linked up with observational sentences.

As an example, consider the example of classical mechanics. Similar to how Euclidean geometry can be fully derived from a set of five axioms, classical mechanics is fully determined by Newton's laws of mechanics. At a purely formal level, it is possible to provide a fully syntactic axiomatization in terms of the relevant symbols, variables, and rules for their manipulation – that is, in terms of what Rudolf Carnap calls the *calculus of mechanics*. If one takes the latter as one's starting point, it requires interpretation of the results derived from within this formal framework, in order for the calculus to be recognizable as a theory of mechanics, that is, of physical phenomena. In the case of mechanics, we may have no difficulty stating the axioms in the form of the (physically interpreted) *Newtonian laws of mechanics*, but in other cases – perhaps in quantum mechanics – making this connection with observables may not be so straightforward. As *Carnap* notes [1.24, p. 57]:

> "[t]he relation of this theory [= the physically interpreted theory of mechanics] to the calculus of mechanics is entirely analogous to the relation of physical to mathematical geometry. "

As in the Euclidean case, the syntactic view identifies the theory with a formal language or calculus (including, in the case of scientific theories, relevant correspondence rules), "whose interpretation – what the calculus is a theory *of* – is fixed at the point of application" [1.25, p. 125].

On the syntactic view of theories, models play at best a very marginal role as limiting cases or approximations. This is for two reasons. First, since the nonobservational part of the theory – that is, the *theory proper*, as one might put it – does not admit of direct interpretation, the route to constructing theoretical models on the basis of our directly interpreting the core ingredients of the theory is obstructed. Interpretation at the level of observational statements, while still available to us, is insufficient to imbue models with anything other than a purely *one-off* auxiliary role. Second, as *Cartwright* has pointedly argued in criticism directed at both the syntactic and the semantic views, there is a shared – mistaken – assumption that theories are a bit like vending machines [1.26, p. 247]:

> "[Y]ou feed it input in certain prescribed forms for the desired output; it gurgitates for a while; then it drops out the sought-for-representation, plonk, on the tray, fully formed, as Athena from the brain of Zeus."

This limits what we can do with models, in that there are only two stages [1.26, p. 247]:

> "First, eyeballing the phenomenon, measuring it up, trying to see what can be abstracted from it that has the right form and combination that the vending machine can take as input; secondly, [. . .] we do either tedious deduction or clever approximation to get a facsimile of the output the vending machine would produce."

Even if this caricature seems a little too extreme, the fact remains that, by modeling theories after first-order formal languages, the syntactic view limits our understanding of what theories and models are and what we can do with them.

1.4.3 The Semantic View

One standard criticism of the syntactic view is that it conflates scientific theories with their linguistic formulations. Proponents of the semantic view argue that by adding a layer of (nonlinguistic) structures between the linguistic formulations of theories and our assessment of them, one can side-step many of the problems faced by the syntactic view. According to the semantic view, a theory should be thought of as the set of set-theoretic structures that satisfy the different linguistic formulations of the theory. A structure that provides an interpretation for, and makes true, the set of sentences associated with a specific linguistic formulation of the theory is called a *model of the theory*. Hence, the semantic view is often characterized as conceiving of theories as *collections of models*. This not only puts models – where these are to be understood in the logical sense outlined earlier – center stage in our account of scientific theories, but also renders the latter fundamentally *extra-linguistic* entities.

An apt characterization of the semantic view is given by *Suppe* as follows [1.27, pp. 82–83]:

> "This suggests that theories be construed as propounded abstract *structures* serving as models for sets of interpreted sentences that constitute the linguistic formulations. [. . .] [W]hat the theory does is directly describe the behavior of abstract systems, known as *physical systems*, whose behaviors depend only on the selected parameters. However, physical systems are abstract replicas of actual phenomena, being what the phenomena *would have been* if no other parameters exerted an influence."

According to a much-quoted remark by one of the main early proponents of the semantic view, *Suppes*, "the meaning of the concept of model is the same in mathematics and in the empirical sciences." However, as *Suppe*'s quote above makes clear, models in science have additional roles to play, and it is perhaps worth noting that *Suppes* himself immediately continues: "The difference to be found in these disciplines is to be found in their use of the concept" [1.28, p. 289]. Supporters of the semantic view often claim that it is closer to the scientific practices of modeling and theorizing than the syntactic view. On this view, according to *van Fraassen* [1.29, p. 64],

> "[t]o present a theory is to specify a family of structures, its *models*; and secondly, to specify certain parts of those models (the *empirical substructures*) as candidates for the direct representation of observable phenomena."

Unlike what the syntactic view suggests, scientists do not typically formulate abstract theoretical axioms and only interpret them at the point of their application to observable phenomena; rather, "scientists build in their mind's eye systems of abstract objects whose properties or behavior satisfy certain constraint (including law)" [1.23, p. 154] – that is, they engage in the construction of theoretical models.

Unlike the syntactic view, then, the semantic view appears to give a more definite answer to the question *what is a model?* In line with the account sketched so far, *a model of a theory is simply a (typically extra-*

linguistic) structure that provides an interpretation for, and makes true, the set of axioms associated with the theory (assuming that the theory is axiomatizable). Yet it is not clear that, in applying their view to actual scientific theories, the semanticists always heed their own advice to treat models as both *giving an interpretation*, and *ensuring the truth*, of a set of statements. More importantly, the model-theoretic account demands that, in a manner of speaking, a model should fulfil its truth-making function *in virtue of* providing an interpretation for a set of sentences. Other ways of ensuring truth – for example by limiting the domain of discourse for a set of fully interpreted sentences, thereby ensuring that the latter will happen to be true – should not qualify. Yet, as *Thomson-Jones* [1.30] has argued, purported applications of the semantic view often stray from the original model-theoretic motivation. As an example, consider Suppes' *axiomatization* of Newtonian particle physics. (The rest of this subsection follows [1.30, pp. 530–531].) *Suppes* [1.31] begins with the following definition (in slightly modified form)

Definition 1.1

A system $\beta = \langle P, T, s, m, f, g\rangle$ is a model of particle mechanics if and only if the following seven axioms are satisfied:

Kinematical axioms:

1 The set P is finite and nonempty
2 The set T is an interval of real numbers
3 For p in P, s_p is twice differentiable.

Dynamical axioms:

4 For p in P, $m(p)$ is a positive real number
5 For p and q in P and t in T,

$$f(p,q,t) = -f(q,p,t)\ .$$

6 For p and q in P and t in T,

$$s(p,t) \times f(p,q,t) = -s(q,t) \times f(q,p,t)\ .$$

7 For p in P and t in T,

$$m(p)D^2 s_p(t) = \sum_{q \in P} f(p,q,t) + g(p,t)\ .$$

At first sight, this presentation adheres to core ideas that motivate the semantic view. It sets out to define an extra-linguistic entity, β, in terms of a set-theoretical predicate; the entities to which the predicate applies are then to be singled out on the basis of the seven axioms. But as *Thomson-Jones* points out, a specific model S defined in this way "is not a serious interpreter of the predicate or the *axioms* that compose it" [1.30, p. 531]; it merely fits a structure to the description provided by the fully interpreted axioms (1)–(7), and in this way ensures that they are satisfied, but it does not make them come out true in virtue of providing an interpretation (i. e., by invoking semantic theory). To *Thomson-Jones*, this suggests that identifying scientific models with truth-making structures in the model-theoretic sense may, at least in the sciences, be an unfulfilled promise of the semantic view; instead, he argues, we should settle for a less ambitious (but still informative) definition of a model as "a mathematical structure used to represent a (type of) system under study" [1.30, p. 525].

1.4.4 Partial Structures

Part of the motivation for the semantic view was its perceived greater ability to account for how scientists actually go about developing models and theories. Even so, critics have claimed that the semantic view is unable to accommodate the great diversity of scientific models and faces special challenges from, for example, the use of inconsistency in many models. In response to such criticisms, a philosophical research program has emerged over the past two decades, which seeks to establish a *middle ground* between the classical semantic view of models discussed in the previous section and those who are sceptical about the prospects of formal approaches altogether. This research program is often called the *partial structures approach*, which was pioneered by *Newton da Costa* and *Steven French* and whose vocal proponents include Otávio Bueno, James Ladyman, and others; see [1.32] and references therein.

Like many adherents of the semantic view, partial structures theorists hold that models are to be reconstructed in set-theoretic terms, as ordered n-tuples of sets: a set of objects with (sets of) properties, quantities and relations, and functions defined over the quantities. A *partial structure* may then be defined as $\mathfrak{A} = \langle D, R_i\rangle_{i\in \mathrm{I}}$, where D is a nonempty set of n-tuples of just this kind and each R_i is a n-ary relation. Unlike on the traditional semantic view, the relations R_i need not be complete isomorphisms, but crucially are *partial relations*: that is, they need not be defined for all n-tuples of elements of D. More specifically, for each partial relation R_i, in addition to the set of n-tuples for which the relation holds and the set of n-tuples for which it does not hold, there is also a third set of n-tuples for which it is underdetermined whether or not it holds. (There is a clear parallel here with Hesse's notion of positive, negative, and neutral analogies which, as *da Costa* and *French* put it, "finds a natural home in the context of partial structures" [1.32, p. 48].) A total structure is said to *extend* a partial structure, if it subsumes the first two sets without change (i. e., includes all those objects and definite relations that exist in the partial structures) and renders each extended relation well defined for every n-tuple of objects in its domain. This gives rise to a hierarchy of structures and substructures, which together with the notion of partial isomorphism loosens the requirements on representation, since all that is needed for two partial models A and A' to be *partially* isomorphic is that a partial substructure of A be isomorphic to a partial substructure in A'.

Proponents of the partial structures approach claim that it "widens the framework of the model-theoretic approach and allows various features of models and theories – such as analogies, iconic models, and so on – to be represented," [1.33, p. 306] that it can successfully contain the difficulties arising from inconsistencies in models, and that it is able to capture "the existence of a hierarchy of models stretching from the data up to the level of theory" [1.33]. Some critics have voiced criticism about such sweeping claims. One frequent criticism concerns the proliferation of partial isomorphisms, many of which will trivially obtain; however, if partial relations are so easy to come by, how can one tell the interesting from the vast majority of irrelevant ones? (*Pincock* speaks in this connection of the "danger of trivializing our representational relationships" [1.34, p. 1254].) *Suárez* and *Cartwright* add further urgency to this criticism, by noting that the focus on set-theoretical structures obliterates all those uses of models and aspects of scientific practice that do not amount to the making of claims [1.35, p. 72]:

> "So all of scientific practice that does not consist in the making of claims gets left out. [...] Again, we maintain that this inevitably leaves out a great deal of the very scientific practice that we are interested in."

It is perhaps an indication of the limitations of the partial structures approach that, in response to such criticism, its proponents need to again invoke heuristic factors, which cannot themselves be subsumed under the proposed formal framework of models as set-theoretic structures with partial relations.

1.5 The Folk Ontology of Models

If we accept that scientific models are best thought of as functional entities (Sect. 1.2), perhaps something can be learnt about the ontology of scientific models from looking at their functional role in scientific inquiry. What one finds across a range of different kinds of models is the practice of taking models as stand-ins for systems that are not, in fact, instantiated. As *Godfrey-Smith* puts it, "modelers often *take* themselves to be describing imaginary biological populations, imaginary neural networks, or imaginary economies" [1.36, p. 735] – that is, they are aware that due to idealization and abstraction, model systems will differ in their descriptions from a full account of the actual world. A model, thus understood, may be thought of as a "description of a missing system," and the corresponding research practice of describing and characterizing model systems *as though* they were real instantiated systems (even though they are not) may be called, following *Thomson-Jones*, the "face-value practice" of scientific modeling [1.37, pp. 285–286].

On the heels of the face-value practice of scientific modeling, it has been argued, comes a common – though perhaps not universally shared – understanding of *what models are* [1.36, p. 735]:

> "[...] to use a phrase suggested by Deena Skolnick, the treatment of model systems as comprising imagined concrete things is the *folk ontology* of at least many scientific modelers. It is the ontology embodied in many scientists' unreflective habits of talking about the objects of their study-talk about what a certain kind of population will do, about whether a certain kind of market will clear. [...O]ne kind of understanding of model-based science requires that we take this *folk ontology* seriously, as part of the scientific strategy."

The ontology of *imagined concrete things* – that is, of entities that, *if real*, would be on a par with concrete objects in the actual world – leads quickly into the thorny territory of fictionalism. *Godfrey-Smith* is explicit about this when he likens models to "something we are all familiar with, the imagined objects of literary fiction" [1.36] – such as Sherlock Holmes, J.R.R. Tolkien's Middle Earth, and so on. Implicit in this suggestion is, of course, a partial answer to our question *What is a model?* – namely, that the ontological status of scientific models is *just like* that of literary (or other) fictions. The advantages and disadvantages of such a position will be discussed in detail in Sect. 1.6 of this chapter.

There is, however, another direction into which a closer analysis of the face-value practice can take us. Instead of focusing on the ontological status of the entities we are imagining when we contemplate models as imagined concrete things, we can focus on the conscious processes that attend such imaginings (or, if one prefers a different way of putting it, the *phenomenology* of interacting with models). Foremost among these is the mental imagery that is conjured up by the descriptions of models. (Indeed, as we shall see in the next section, on certain versions of the fictionalist view, a model *prescribes* imaginings about its target system.) How much significance one should attach to the mental pictures that attend our conscious consideration of models has been a matter of much controversy: recall Duhem's dismissal of mechanical imagery as a way of conceptualizing electromagnetic phenomena (Sect. 1.3).

Focusing on the mental processes that accompany the use of scientific models might lead one to propose an analysis of models in terms of their cognitive foundations. Nancy Nersessian has developed just such an analysis, which ties the notion of models in science closely to the cognitive processes involved in mental modeling. Whereas the traditional approach in psychology had been to think of reasoning as consisting of the mental application of logical rules to propositional representations, mounting empirical evidence of the role of heuristics and biases suggested that much of human reasoning proceeds via *mental models* [1.38], that is, by carrying out thought experiments on internal models. A *mental model*, on this account, is "a structural analog of a real-world or imaginary situation, event, or process" as constructed by the mind in reasoning (and, presumably, realized by certain underlying brain processes) [1.39, pp. 11–12]:

> "What it means for a mental model to be a structural analog is that it embodies a representation of the spatial and temporal relations among, and the causal structures connecting the events and entities depicted and whatever other information that is relevant to the problem-solving talks. [...] The essential points are that a mental model can be nonlinguistic in form and the mental mechanisms are such that they can satisfy the model-building and simulative constraints necessary for the activity of mental modeling."

While this characterization of mental models may have an air of circularity, in that it essentially defines mental models as place-holders for *whatever it takes* to support *the activity of mental modeling*, it nonetheless suggests a place to look for the materials from which models are constructed: the mind itself, with its various types of content and mental representation. As *Nersessian* puts it: "Whatever the format of the model itself, information in various formats, including linguistic, formulaic, visual, auditory, kinesthetic, can be used in its construction" [1.39, p. 12].

How does this apply to the case of *scientific* models? As an example, Nersessian considers James Clerk Maxwell's famous molecular vortex model, which visualized the lines of magnetic force around a magnet as though they were vortices within a continuous fluid (Fig. 1.1).

As Nersessian sees it, Maxwell's drawing "is a *visual* representation of an *analogical* model that is accompanied with instructions for *animating* it correctly in thought" [1.39, p. 13]. And indeed *Maxwell* gives detailed instructions regarding how to interpret, and bring to life, the model of which the reader is only given a momentary *snapshot* [1.40, p. 477]:

> "Let the current from left to right commence in *AB*. The row of vortices *gh* above *AB* will be set in motion in the opposite direction to a watch [...]. We shall suppose the two of vortices *kl* still at rest, then the layer of particles between these rows will be acted on by the row *gh*,"

and so forth. It does seem plausible to say that such instructions are intended to prescribe certain mental models on the part of the reader. Convincing though this example may be, it still begs the question of what, *in general*, a mental model is. At the same time, it illustrates what is involved in conjuring up a mental model and which materials – in this case, spatial representations, along with intuitions about the mechanical motion of parts in a larger system – are involved in its constitution.

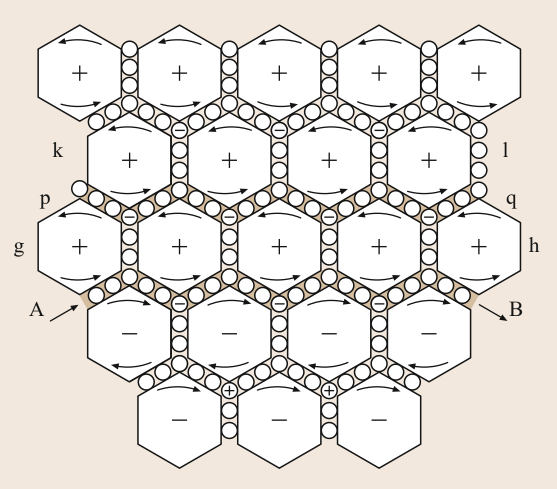

Fig. 1.1 Maxwell's drawing of the molecular vortex model (after [1.40])

1.6 Models and Fiction

As noted in the previous section, the face-value practice of scientific modeling and its concomitant folk ontology, according to which models are imagined concrete things, have a natural affinity to the way we think about fictions. As one proponent of models as fictions puts it [1.41, p. 253]:

> "The view of model systems that I advocate regards them as imagined physical systems, that is, as hypothetical entities that, as a matter of fact, do not exist spatiotemporally but are nevertheless not purely mathematical or structural in that they would be physical things if they were real."

Plausible though this may sound, the devil is in the details. A first – perhaps trivial – caveat concerns the restriction that model systems *would be physical things if they were real*. In order to allow for the notion of model to be properly applied to the social and cognitive sciences, such as economics and psychology, it is best to drop this restriction to physical systems. (On this point, see [1.30, p. 528].) This leaves as the gist of the folk-ontological view the thought that model systems, *if* they were real, would be *just as we imagine them* (or, more carefully, *just as the model instructs us to imagine them*).

In order to sharpen our intuitions about fictions, let us introduce an example of a literary fiction, such as the following statement from *Doyle*'s *The Adventure of the Three Garridebs* (1924) [1.42]: "Holmes had lit his pipe, and he sat for some time with a curious smile upon his face." There is, of course, no actual human being that this statement represents: no one is sitting smilingly at 221B Baker Street, filling up the room with smoke from their pipe. (Indeed, until the 1930s, the address itself had no real-world referent, as the highest number on Baker Street then was No. 85.) And yet there is a sense in which this passage does seem to represent Sherlock Holmes and, within the context of the story, tells us something informative about him. In particular, it seems to lend support to certain statements about Sherlock Holmes as opposed to others. If we say *Holmes is a pipe smoker*, we seem to be asserting something true about him, whereas if we say *Holmes is a nonsmoker*, we appear to be asserting something false. One goal of the ontology of fictions is to make sense of this puzzle.

Broadly speaking, there are two kinds of philosophical approaches – realist and antirealist – regarding fictions. On the realist approach, even though Sherlock Holmes is not an actual human being, we must grant that he *does* exist in some sense. Following *Meinong* [1.43], we might, for example, distinguish between *being* and *existence* and consider Sherlock Holmes to be an object that has all the requisite properties we normally attribute to him, except for the property of existence. Or we might take fictions to have existence, but only as abstract entities, not as objects in space and time. By contrast, antirealists about fictions deny that they have independent being or existence and instead settle for other ways of making sense of how we interpret fictional discourse. Following Bertrand Russell, we might paraphrase the statement *Sherlock Holmes is a pipe smoker and resides at 221B Baker Street* without the use of a singular term (*Sherlock Holmes*), solely in terms of a suitably quantified existence claim: *There exists one and only one x such that x is a pipe smoker and x resides at 221B Baker Street*. However, while this might allow us to parse the meaning of further statements about Sherlock Holmes more effectively, it does not address the puzzle that certain claims (such as *He is a pipe smoker*) ring true, whereas others do not – since it renders each part of the explicated statement false. This might not seem like a major worry for the case of literary fictions, but it casts doubt on whether we can fruitfully think about scientific models in those terms, given the epistemic role of scientific models as contributors to scientific knowledge.

In recent years, an alternative approach to fictions has garnered the attention of philosophers of science, which takes *Walton*'s notion of "games of make-believe" as its starting point. *Walton* introduces this notion in the context of his philosophy of art, where he characterizes (artistic) representations as "things possessing the social function of serving as props in games of make-believe" [1.44, p. 69]. In games of make-believe, participants engage in behavior akin to children's pretend play: when a child uses a banana as a telephone *to call grandpa*, this action does not amount to actually calling her grandfather (and perhaps not even *attempting* to call him); rather, it is a move within the context of play – where the usual standards of realism are suspended – whereby the child resolves to treat the situation *as if* it were one of speaking to her grandfather on the phone.

The banana is simply a prop in this game of make-believe. The use of the banana as a make-believe telephone may be inspired by some physical similarity between the two objects (e.g., their elongated shape, or the way that each can be conveniently held to one's ear and mouth at the same time), but it is clear that props can go beyond material objects to include, for example, linguistic representations (as would be the case with the literary figure of Sherlock Holmes). While the rules governing individual pretend play may be ad hoc, communal games of make-believe are structured by shared normative principles which *authorize* certain moves as legitimate, while excluding other moves as illegitimate. It is in virtue of such principles that fictional truths can be generated: for example, a toy model of a bridge at the scale of 1 : 1000 prescribes that, "if part of the model has a certain length, then, fictionally, the corresponding part of the bridge is a thousand times that length" [1.45, p. 38] – in other words, even though the model itself is only a meter long, it *represents* the bridge *as* a thousand meters long. Note that the scale model could be a model of a bridge that is yet to be built – in which case it would still be true that, fictionally, the bridge is a thousand meters long: props, via the rules that govern them, *create* fictional truths.

One issue of contention has been what kinds of metaphysical commitments such a view of models entails. Talk of *imagined concrete things* as the material from which models are built has been criticized for amounting to an indirect account of modeling, by which [1.46, pp. 308, fn. 14]

> "prepared descriptions and equations of motion ask us to imagine an *imagined concrete system* which then bears some other form of representation relation to the system being modelled."

A more thoroughgoing direct view of models as fictions is put forward by *Toon*, who considers the following sentence from *Wells*'s *The War of the Worlds*: "The dome of St. Paul's was dark against the sunrise, and injured, I saw for the first time, by a huge gaping cavity on its western side" [1.47, p. 229]. As *Toon* argues [1.46, p. 307]:

> "There is no pressure on us to postulate a fictional, damaged, St. Paul's for this passage to represent; the passage simply represents the actual St. Paul's. Similarly, on my account, our prepared description and equation of motion do not give rise to a fictional, idealised bouncing spring since they represent the actual bouncing spring."

By treating models as prescribing imaginings about *the actual objects* (where these exist and are the model's target system), we may resolve to imagine all sorts of things that are, as a matter of fact, false; however, so the direct view holds, this is nonetheless preferable to the alternative option of positing *independently existing* fictional entities [1.45, p. 42]. Why might one be tempted to posit, as the indirect view does, that fictional objects fitting the model descriptions must exist? An important motivation has to do with the assertoric force of our model-based claims. As *Giere* puts it: "If we insist on regarding principles as genuine statements, we have to find something that they describe, something to which they refer" [1.48, p. 745]. In response, proponents of the direct view have disputed the need "to regard theoretical principles formulated in modeling as genuine statements"; instead, as *Toon* puts it, "they are prescriptions to imagine" [1.45, p. 44].

One potential criticism the models as fictions view needs to address is the worry that, by focusing on the user's imaginings, what a model is becomes an entirely subjective matter. A similar worry may be raised with respect to the mental models view discussed in Sect. 1.5: if a model is merely a place-holder for whatever is needed to sustain the activity of mental modeling (or imagining) on the part of an agent, how can one be certain that the same kinds of models (or props) reliably give rise to the same kinds of mental modeling (or imaginings)? In this respect, at least, the models as fictions view appears to be in a stronger position. Recall that, unlike in individual pretend play (or unconstrained imagining), in games of make-believe certain imaginations are sanctioned by the prop itself and the – public, shared – rules of the game. As a result, "someone's imaginings are governed by intersubjective rules, which guarantee that, as long as the rules are respected, everybody involved in the game has the same imaginings" [1.41, p. 264] – though it should be added, not necessarily the same *mental images*.

In his 1963 book, *Models and Metaphors*, *Black* expressed his hope that an "exercise of the imagination, with all its promise and its dangers" may help pave the way for an "understanding of scientific models and archetypes" as "a reputable part of scientific culture" [1.4, p. 243]. Even though Black was writing in general terms (and perhaps for rhetorical effect), his characterization would surely be considered apt by the proponents of the models as fictions view, who believe that models allow us to imagine their targets to be a certain way, and that, by engaging in such imaginings, we can gain new scientific insights.

1.7 Mixed Ontologies: Models as Mediators and Epistemic Artifacts

In Sect. 1.1, a distinction was drawn between *informational* views of models, which emphasize the objective, two-place relation between the model and what it represents, and *pragmatic* views, according to which a model depends at least in part on the user's beliefs or intentions, thereby rendering model-based representation a three-place relation between model, target, and user. Unsurprisingly, which side one comes down on in this debate will also have an effect on one's take on the ontology of scientific models. Hence, structuralist approaches (e.g., the partial structures approach discussed in Sect. 1.4.4) are a direct manifestation of the informational view, whereas the models as fictions approach – especially insofar as it considers models to be props for the user's imagination – would be a good example of the pragmatic view. The pragmatic dimension of scientific representation has received growing attention in the philosophical literature, and while this is not the place for a detailed survey of pragmatic accounts of model-based representation in particular, the remainder of this section will be devoted to a discussion of the ontological consequences of several alternative pragmatic accounts of models. Particular emphasis will be placed on what I shall call *mixed ontologies*, that is, accounts of models that emphasize the heterogeneity and diversity of their components.

1.7.1 Models as Mediators

Proponents of pragmatic accounts of models usually take scientific practice as the starting point of their analysis. This often directly informs how they think about models; in particular, it predisposes them to treat models as the outcome of a process of model construction. On this view, it is not only the *function* of models – for example, their capacity to represent target systems – which depends on the beliefs, intentions, and cognitive interests of a model user, but also the very *nature* of models which is dependent on human agents in this way. In other words, what models are is crucially determined by their being the result of a deliberate process of model construction. Model construction, most pragmatic theorists of models insist, is marked by "piecemeal borrowing" [1.35, p. 63] from a range of different domains. Such conjoining of heterogeneous components to form a model cannot easily be accommodated by structuralist accounts, or so it has been claimed; at the very least, there is considerable tension between, say, the way that the partial structures approach allows for a nested *hierarchy* of models (connected with one another via partial isomorphisms) and the much more ad hoc manner in which modelers piece together models from a variety of ingredients. (On this point, see especially [1.35, p. 76].)

A number of such accounts have coalesced into what has come to be called the *models as mediators* view (see [1.49] for a collection of case studies). According to this view, models are to be regarded neither as a merely auxiliary intermediate step in applying or interpreting scientific theories, nor as constructed purely from data. Rather, they are thought of as mediating between our theories and the world in a partly autonomous manner. As *Morrison* and *Morgan* put it, models "are *not* situated in the middle of an hierarchical structure between theory and the world," but operate outside the hierarchical "theory-world axis" [1.50, pp. 17–18]. A central tenet of the models as mediators view is the thesis that models "are made up from a *mixture* of elements, including those from outside the domain of investigation"; indeed, it is thought to be precisely in virtue of this heterogeneity that they are able to retain "an element of independence from both theory and data (or phenomena)" [1.50, p. 23].

At one level, the models as mediators view appears to be making a descriptive point about scientific practice. As *Morrison* and *Morgan* [1.50] point out, there is "no *logical* reason why models should be constructed to have these qualities of partial independence" [1.50, p. 17], though in practice they do exhibit them, and examples that involve the integration of heterogeneous elements beyond theory and data "are not the exception but the rule" [1.50, p. 15]. Yet, there is also the further claim that models could not fulfil their epistemic function *unless* they are partially autonomous entities: "we can only expect to use models to learn about our theories or our world if there is at least partial independence of the model from both" [1.50, p. 17]. Given that models are functional entities (in the sense discussed in Sect. 1.2), this has repercussions for the ontological question of what kind of entities models are. More often than not, models will integrate – perhaps imperfectly, but in irreducible ways – heterogeneous components from disparate sources, including (but not limited to) "elements of theories and empirical evidence, as well as stories and objects which could form the basis for modeling decisions" [1.50, p. 15]. As proponents of the models as mediators view are at pains to show, even in cases where models initially seem to derive straightforwardly from fundamental theory or empirical data, closer inspection reveals the presence of other elements – such as "simplifications and approximations which have to be decided independently of the theoretical requirements or of data conditions" [1.50, p. 16].

For the models as mediators approach, any answer to the question *what is a model?* must be tailored to the specific case at hand: models in high-energy physics will have a very different composition, and will consist of an admixture of different elements, than, say, models in psychology. However, as a general rule, no model – or, at any rate, no *interesting* model – will ever be fully reducible to theory and data; attempts to *clean up* the ontology of scientific models so as to render them either purely theoretical or entirely empirical, according to the models as mediators view, misconstrue the very nature and function of models in science.

1.7.2 Models as Epistemic Artifacts

A number of recent pragmatic approaches take the models as mediators view as their starting point, but suggest that it should be extended in various ways. Thus, *Knuuttila* acknowledges the importance of mediation between theory and data, but a richer account of models is needed to account for how this partial independence comes about. For *Knuuttila*, *materiality* is the key enabling factor that imbues models with such autonomy: it is "the material dimension, and not just *additional elements*, that makes models able to mediate" [1.51, p. 48]. Materiality is also seen as explaining the various epistemic functions that models have in inquiry, not least by way of analogy with scientific experiments. For example, just as in experimentation much effort is devoted to minimizing unwanted external factors (such as noise), in scientific models certain methods of approximation and idealization serve the purpose of neutralizing undesirable influences. Models typically draw on variety of formats and representations, in a way that *enables* certain specific uses, but at the same time *constrains* them; this breaks with the traditional assumption that we can "clearly tell apart those features of our scientific representations that are attributable to the phenomena described from the conventions used to describe them" [1.52, p. 268].

On the account sketched thus far, attempting to characterize the nature and function of models in the language of theories and data would, in the vast majority of cases, give a misleading impression; instead, models are seen as *epistemic tools* [1.52, p. 267]:

> "Concrete artifacts, which are built by various representational means, and are constrained by their design in such a way that they enable the study of certain scientific questions and learning through constructing and manipulating them."

This links the philosophical debate about models to questions in the philosophy of technology, for example concerning the ontology of artifacts, which are likewise construed as both material bodies and functional objects. It also highlights the constitutive role of design and construction, which applies equally to models with a salient material dimension – such as scale models in engineering or ball-and-stick models in chemistry – and to largely theoretical models. For example, it has been argued that mathematical models (e.g., in many-body physics) may be fruitfully characterized not only in theoretical terms (say, as a Hamiltonian) or as mathematical entities (as an operator equation), but also as the output of a *mature mathematical formalism* (in this case, the formalism of second quantization) – that is, a physically interpreted set of notational rules that, while embodying various theoretical assumptions, is not usually reducible to fundamental theory [1.53].

As in the case of the models as mediators approach, the ontological picture that emerges from the artifactual approach to models is decidedly mixed: models will typically consist of a combination of different materials, media and formats, and deploy different representational means (such as pictorial, symbolic, and diagrammatic notations) as well as empirical data and theoretical assumptions. Beyond merely acknowledging the heterogeneity of such a *mixture of elements*, however, the artifactual approach insists that it is *in virtue of their material dimension* that the various elements of a model, taken together, enable and constrain its representational and other epistemic functions.

1.8 Summary

As the survey in this chapter demonstrates, the term *model* in science refers to a great variety of things: physical objects such as scale models in engineering, descriptions and sets of sentences, set-theoretic structures, fictional objects, or an assortment of all of the above. This makes it difficult to arrive at a uniform characterization of models *in general*. However, by paying close attention to philosophical accounts of model-based representation, it is possible to discern certain clusters of positions. At a general level, it is useful to think of models as functional entities, as this allows one to explore how different functional perspectives lead to different conceptions of the ontology of models. Hence, with respect to the representational function of models, it is possible to distinguish between *informational* views, which we found to be closely associated with structuralist accounts of models, and *pragmatic* views, which tend to give rise to more heterogeneous accounts, according to which models may be thought of as *props for the imagination*, as partly autonomous mediators between theory and data, or as epistemic artifacts consisting of an admixture of heterogeneous elements.

When nineteenth century physicists began to reflect systematically on the role of *analogy* in science, they did so out of a realization that it may not always be possible to apply fundamental theory directly to reality, either because any attempt to do so faces insurmountable complexities, or because no such fundamental theory is as yet available. At the beginning of the twenty-first century, these challenges have not diminished, and scientists find themselves turning to an ever greater diversity of scientific models, a unified philosophical theory of which is still outstanding.

References

1.1 R. Frigg: Models in science. In: *Stanford Encyclopedia of Philosophy*, ed. by E.N. Zalta http://plato.stanford.edu/entries/models-science/ (Spring 2012 Edition)

1.2 N. Goodman: *Languages of Art* (Bobbs-Merrill, Indianapolis 1968)

1.3 R. Ankeny, S. Leonelli: What's so special about model organisms?, Stud. Hist. Philos. Sci. **42**(2), 313–323 (2011)

1.4 M. Black: *Models and Metaphors: Studies in Language and Philosophy* (Cornell Univ. Press, Ithaca 1962)

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

1.6 J. von Neumann: Method in the physical sciences. In: *Collected Works Vol. VI. Theory of Games, Astrophysics, Hydrodynamics and Meteorology*, ed. by A.H. Taub (Pergamon, Oxford 1961) pp. 491–498

1.7 S. French: Keeping quiet on the ontology of models, Synthese **172**(2), 231–249 (2010)

1.8 G. Contessa: Editorial introduction to special issue, Synthese **2010**(2), 193–195 (2010)

1.9 S. Ducheyne: Towards an ontology of scientific models, Metaphysica **9**(1), 119–127 (2008)

1.10 R. Giere: Using models to represent reality. In: *Model-Based Reasoning in Scientific Discovery*, ed. by L. Magnani, N. Nersessian, P. Thagard (Plenum, New York 1999) pp. 41–57

1.11 A. Chakravartty: Informational versus functional theories of scientific representation, Synthese **217**(2), 197–213 (2010)

1.12 R.I.G. Hughes: Models and representation, Proc. Philos. Sci., Vol. 64 (1997) pp. S325–226

1.13 T. Knuuttila: Some consequences of the pragmatist approach to representation. In: *EPSA Epistemology and Methodology of Science*, ed. by M. Suárez, M. Dorato, M. Rédei (Springer, Dordrecht 2010) pp. 139–148

1.14 M. Suárez: An inferential conception of scientific representation, Proc. Philosophy of Science, Vol. 71 (2004) pp. 67–779

1.15 M. Jammer: Die Entwicklung des Modellbegriffs in den physikalischen Wissenschaften, Stud. Gen. **18**(3), 166–173 (1965)

1.16 P. Duhem: *The Aim and Structure of Physical Theory* (Princeton Univ. Press, Princeton 1954), Transl. by P.P. Wiener

1.17 D. Bailer-Jones: Models, metaphors and analogies. In: *The Blackwell Guide to the Philosophy of Science*, ed. by P. Machamer, M. Silberstein (Blackwell, Oxford 2002) pp. 108–127

1.18 D.H. Mellor: Models and analogies in science: Duhem versus Campbell?, Isis **59**(3), 282–290 (1968)

1.19 D. Bailer-Jones: *Scientific Models in Philosophy of Science* (Univ. Pittsburgh Press, Pittsburgh 2009)

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

1.21 N.R. Campbell: *Physics: The Elements* (Cambridge Univ. Press, Cambridge 1920)

1.22 J.M. Soskice, R. Harré: Metaphor in science. In: *From a Metaphorical Point of View: A Multidisciplinary Approach to the Cognitive Content of Metaphor*, ed. by Z. Radman (de Gruyter, Berlin 1995) pp. 289–308

1.23 C. Liu: Models and theories I: The semantic view revisited, Int. Stud. Philos. Sci. **11**(2), 147–164 (1997)

1.24 R. Carnap: *Foundations of Logic and Mathematics* (Univ. Chicago Press, Chicago 1939)

1.25 R. Hendry, S. Psillos: How to do things with theories: An interactive view of language and models in science. In: *The Courage of Doing Philosophy: Essays Presented to Leszek Nowak*, ed. by J. Brzeziński, A. Klawiter, T.A.F. Kuipers, K. Lastowksi, K. Paprzycka, P. Przybyzs (Rodopi, Amsterdam 2007) pp. 123–158

1.26 N. Cartwright: Models and the limits of theory: Quantum hamiltonians and the BCS model of superconductivity. In: *Models as Mediators*, ed. by M. Morrison, M. Morgan (Cambridge Univ. Press, Cambridge 1999) pp. 241–281

1.27 F. Suppe: *The Semantic Conception of Theories and Scientific Realism* (Univ. Illinois Press, Urbana 1989)

1.28 P. Suppes: A comparison of the meaning and uses of models in mathematics and the empirical sciences, Synthese **12**(2/3), 287–301 (1960)

1.29 B. van Fraassen: *The Scientific Image* (Oxford Univ. Press, Oxford 1980)

1.30 M. Thomson-Jones: Models and the semantic view, Philos. Sci. **73**(4), 524–535 (2006)

1.31 P. Suppes: *Introduction to Logic* (Van Nostrand, Princeton 1957)

1.32 N. da Costa, S. French: *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning* (Oxford Univ. Press, New York 2003)

1.33 S. French: The structure of theories. In: *The Routledge Companion to Philosophy of Science*, 2nd edn., ed. by M. Curd, S. Psillos (Routledge, London 2013) pp. 301–312

1.34 C. Pincock: Overextending partial structures: Idealization and abstraction, Philos. Sci. **72**(4), 1248–1259 (2005)

1.35 M. Suárez, N. Cartwright: Theories: Tools versus models, Stud. Hist. Philos. Mod. Phys. **39**(1), 62–81 (2008)

1.36 P. Godfrey-Smith: The strategy of model-based science, Biol. Philos. **21**(5), 725–740 (2006)

1.37 M. Thomson-Jones: Missing systems and the face value practice, Synthese **172**(2), 283–299 (2010)

1.38 P.N. Johnson-Laird: *Mental Models* (Harvard Univ. Press, Cambridge 1983)

1.39 N. Nersessian: Model-based reasoning in conceptual change. In: *Model-Based Reasoning in Scientific Discovery*, ed. by L. Magnani, N. Nersessian, P. Thagard (Plenum, New York 1999) pp. 5–22

1.40 J.C. Maxwell: *The Scientific Papers of James Clerk Maxwell*, Vol. 1 (Cambridge Univ. Press, Cambridge 1890), ed. by W.D. Niven

1.41 R. Frigg: Models and fiction, Synthese **172**(2), 251–268 (2010)

1.42 A.C. Doyle: The Adventure of the Three Garridebs. In: *The Casebook of Sherlock Holmes*, ed. by J. Miller (Dover, 2005)

1.43 A. Meinong: *Untersuchungen zur Gegenstandstheorie und Psychologie* (Barth, Leipzig 1904)

1.44 K. Walton: *Mimesis as Make-Believe: On the Foundations of the Representational Arts* (Harvard Univ. Press, Cambridge 1990)

1.45 A. Toon: *Models as Make-Believe: Imagination, Fiction, and Scientific Representation* (Palgrave-Macmillan, Basingstoke 2012)

1.46 A. Toon: The ontology of theoretical modelling: Models as make-believe, Synthese **172**(2), 301–315 (2010)

1.47 H.G. Wells: *War of the Worlds* (Penguin, London 1897), 1978

1.48 R. Giere: How models are used to represent reality, Proc. Philosophy of Science, Vol. 71 (2004) pp. S742–S752

1.49 M. Morrison, M. Morgan (Eds.): *Models as Mediators* (Cambridge Univ. Press, Cambridge 1999)

1.50 M. Morrison, M. Morgan: Models as mediating instruments. In: *Models as Mediators*, ed. by M. Morrison, M. Morgan (Cambridge Univ. Press, Cambridge 1999) pp. 10–37

1.51 T. Knuuttila: *Models as Epistemic Artefacts: Toward a Non-Representationalist Account of Scientific Representation* (Univ. Helsinki, Helsinki 2005)

1.52 T. Knuuttila: Modelling and representing: An artefactual approach to model-based representation, Stud. Hist. Philos. Sci. **42**(2), 262–271 (2011)

1.53 A. Gelfert: *How to Do Science with Models: A Philosophical Primer* (Springer, Cham 2016)