Using Models to Represent Reality

Ronald N. Giere
Department of Philosophy
Center for Philosophy of Science
University of Minnesota

1. Introduction. There has recently been an increase in interest in the role of models in science, of which the Pavia workshop on model-based reasoning is a manifestation. One result of this increased attention has been a proliferation of views on what models are and how they are used in science. In this presentation I will develop a unified interpretation of the nature and role of models in science. Central to this interpretation is an understanding of the relationships between models and other elements of an understanding of science, particularly theories, data, and analogy. My conclusion will be that models play a much larger role in science than even the most ardent enthusiasts for models have typically claimed. Modeling, on my view, is not at all ancillary to doing science, but central to constructing scientific accounts of the natural world.

When I say I seek an *interpretation* of the nature and roles of models in science, I allow that other interpretations are possible. There is no unique essence to the nature of models that might be revealed by philosophical analysis. Nevertheless, I think my interpretation is better than some others, and I will attempt to convince you that this is so.

2. Model Theory. The claim that an understanding of models is central to an understanding of science is not new. Almost forty years ago, Patrick Suppes (1960) published a much cited paper with the title: A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences. The thesis of this paper was that the meaning and use of models can be interpreted as being the *same* in the empirical sciences as it is in mathematics, and, more particularly, in mathematical logic.

At the time Suppes wrote this paper, the theory of models was closely tied with logic. Thus Suppes wrote: A theory is a linguistic entity consisting of a set of sentences and models are non-linguistic entities in which the theory is satisfied. More specifically, a model for Suppes is a set-theoretical structure consisting of a set of objects together with properties, relations, and functions defined over the set of objects. The important point is that when the specific objects, properties, relations, and functions are coordinated with terms in the axioms of a theory, the axioms all come out to be true, relative, of course, to our prior understanding of the domain of objects considered. So a model, on this account, provides an *interpretation* of a set of uninterpreted axioms. For this reason such models are often called “interpretive models.” They might also be called “instantial models” since they instantiate the axioms of a theory understood as comprised of linguistic (including mathematical) statements.

For logicians, most of the models considered consist of *abstract* entities such as numbers or geometrical points and lines. *In principle*, however, the objects considered could be *physical* objects, such as the Earth and the Moon. This is the basis for Suppe's claim that the concept of a model is the same in empirical science as in mathematics. Later I will consider whether any interesting scientific theories, suitably reconstructed, might *in fact* have physical models.

I should note that over the past forty years, as the study of abstract models moved from the province of philosophers and logicians to that of mathematicians, the connection between model theory and logic has become rather attenuated. Current treatises on model theory, such as that by Hodges (1993), focus directly on what are called structures, which are abstract, non-linguistic entities. For example, the groups of group theory and vector spaces are structures in this sense. It would, therefore, be misleading to refer to the models of contemporary mathematical model theory exclusively as interpretive or instantial models.

It is Suppes' view, however, that became, and largely remains, *a*, if not *the*, standard view of models within the philosophy of science. Let me highlight the point that Suppes' view of models incorporates a quite specific relationship between a *theory* (a set of axioms) and a *model* (a set of objects satisfying the axioms).

3. Instantial Models and Analogy. The instantial conception of models supports a quite specific understanding of the nature of *analogy* in science. Uninterpreted logical formulae may be interpreted using many different instantial models. These models will all be *isomorphic*, that is, there will be a one-to-one correspondence between the elements of the different models. This provides a basis for saying that the corresponding elements of the models, as well as the models themselves, are *analogous*. Of course the models in question must be *physical*, and not merely mathematical.

A classic example of such an analogy is that between, on the one hand, an electrical circuit consisting of an inductance coil and a capacitor (an L-C circuit) and, on the other hand, a mechanical oscillator, such as a bouncing spring. Here the inductance coil is said to be analogous to the mass on the oscillating spring while the capacitance is said to be analogous to the spring constant. The current at any point in the circuit is then analogous to the position of the mass on the spring. Both exhibit sinusoidal change over time. The differential equations describing both change in current and change in position have exactly the same abstract form. These relationships are shown in Figure 1.

[Insert Figure 1 here]

4. Representational Models. The instantial conception of models is a well-defined conception of considerable value, particularly in the study of formal logic and the foundations of mathematics. Nevertheless, contrary to Suppes' claims, I do not think it is the best conception of models for understanding models as used in the practice of the empirical sciences. I will not directly criticize the instantial conception of models as a means for understanding scientific practice. Rather, I will simply point out the difficulties I see from my own, alternative, perspective.

For contrast, I call my understanding of models *representational* because it takes models not primarily as providing a means for interpreting formal systems, but as tools for *representing the world*. That is not their only function, but it is, I think, the central function of models used in empirical science. For the moment, then, forget about logic and concentrate on scientific practice, indeed, on the practice of a lowly science, *cartography***.**

Maps. Figure 2 is a standard tourist map of the central area of Pavia. Let us explore some relevant properties of maps. First, maps are not linguistic entities. They are *physical objects*, for example, a piece of paper with lines on it. It does not, therefore, strictly make sense to ask whether a map is *true* or *false*. Those designations are usually reserved for linguistic entities. Moreover, maps are not usually thought of as *instantiations* of any linguistic forms. Of course, one can create a language-like version of any map by creating a digitized bit-map, as was done to produce Figure 2. One could then say that the map of Figure 2 is an instantiation of a long, uninterpreted binary code. But this is an extreme way to secure a conception of maps as instantial models. Such an interpretation plays absolutely no role in understanding the nature or function of maps, which were well-established long before anyone even had the idea of a bit-map. Nevertheless, even though they are neither linguistic entities nor instantiations of linguistic entities, maps are *representational*. Just *how* they are representational is another question, one I will take up shortly. First let us consider some further characteristics of maps.

[Insert Figure 2 here]

Maps are *partial*. Only some features of the territory in question are represented. For example, the map of Figure two represents very few buildings. Moreover, even the indicated features are not fully specified, such as the height of the Palazzo Universita. Maps are of *limited accuracy* regarding included features. Relative distances on the map, for example, will not correspond exactly to relative distances on the ground. This could not be otherwise. No real map could possibly indicate literally all features of a territory with perfect accuracy. In the limit, the only perfect map of a territory would be the territory itself, which would no longer be a map at all. Here one may recall the story by Borges (1954) in which the cartographers of a fictional land set out to construct a map of their land on a scale of one to one. As they complete their project, the people of the land begin to move onto the new territory. The map of Figure 2, by contrast, is a *representational model* of Pavia. It *represents* Pavia in its own special way.

Let us now return to the question: *How* does this map represent Pavia? The answer is: by being *spatially similar to* aspects of Pavia. For example, the lines on the map have similar spatial orientations to some streets in Pavia. In using a map we are using features of one two dimensional surface (the map) to represent features of another two dimensional surface (the surface of the city). For example, some lines on the map represent streets in the city. To generalize: A map represents the region mapped in virtue of shared spatial similarities between the map and the region mapped. Here one *object* (a map) is used to represent another *object* (a geographic region). This notion is explicitly opposed to that of a statement representing a state of affairs.

Similarity versus Isomorphism. Philosophers tend to be suspicious of appeals to similarity. A standard objection is that, since anything is similar to anything else in some respects or other, claims of similarity are vacuous. One might be tempted to invoke isomorphism here, saying that the map is isomorphic, or partially isomorphic, to aspects of the city. But this just cannot be right. No reasonably detailed map can be accurate enough to exhibit a literal isomorphism with identifiable features of a real geographical surface. So the best one could do is invoke something like "approximate isomorphism". Absent some account of what "approximate" might mean in this context, such talk only gives the appearance of clarity. It offers no real conceptual advantages over talk of similarity. And it may disguise problems that need to be faced directly.

Here, by the way, is a basis for questioning the idea that there can *in fact* be *physical* instantiations of the statements of a linguistically formulated theory. It seems easy enough to imagine the objects of an instantial model being physical objects, such as the Earth, Moon, or the planets. But as soon as one adds quantitative functions, such as the mass of the Earth or the distance between the Earth and the Moon, one is in great danger of ending up with false statements, which is to say, no model at all.

Charges of vacuity for claims of similarity can be met by specifying 1) the respects in which the map is said to be similar to the region mapped, and 2) the degree of similarity regarding these respects. Thus, a map might be highly accurate with respect to relative linear distances, but contain very little information about relative elevations. Here an important general point is that the respects and degrees of similarity must be specified from the outside, so to speak. They are not intrinsic to any map or geographical region. Thus maps necessarily reflect the interests of map makers and map users. Maps are *interest relative*, and necessarily so.

Not only have philosophers been suspicious of the concept of similarity, they typically claim there is no way to give a satisfactory *general* account of the notion of similarity. But there is no need to look for a general account of similarity between a model and what is modeled. Similarity is *context dependent*. In any particular context, what is said to be similar to what, in what ways, and to what degrees, can be specified. Of course, there is no unique specification. There are many possible specifications depending on the particular interests of those doing the modeling.

These points may be reinforced by considering a somewhat more abstract kind of map, a subway map, as exhibited in Figure 3. Here spatial locations are indicated only very schematically. The important information is *topological*. One gets the ordering of stations on individual lines together with indications where two lines meet and thus where transfers from one line to another are possible. So the important similarities are those between these topological features of the map and of the whole metro system.

[Insert Figure 3 here]

5. Other Material Models. *Diagrams*. There are many types of diagrams. I will restrict my comments to two dimensional line drawings, such as the circuit diagram shown in Figure 4. The similarity between maps and diagrams is obvious. One could call this a map of the electrical circuit. It shows the pathways electricity can follow. Here I want to say: The diagram is a *representational model* of the circuit. Again we have *one thing*, the diagram on paper, being used to represent *another thing*, an electrical circuit.

[Insert Figure 4 here]

In this diagram, the spatial locations of the wires are not important. There need be no strong similarity between the relative positions of the wires in the diagram and in the physical circuit. What matters is only what is connected to what. So what is modeled are connections, not spatial locations. Connections are more abstract than locations. The locations of lines representing wires in the diagram should be organized so as to make it easy for the human eye and brain to perceive the connections. How things are actually wired is a matter of convenience or efficiency in the physical wiring process.

Scale Models. There are many sorts of scale models, from model houses to models of the solar system. A canonical example in twentieth century science is the three dimensional scale model Jim Watson built in the process of discovering the double helical structure of DNA molecules. This scale model was a representational model of DNA molecules. And it was representational in virtue of three-dimensional spatial and structural similarities between the scale model and real DNA molecules. The base pairs of DNA were claimed to be arranged in a helical structure similar to the pieces of tin and cardboard in Watson's scale model. Here again we have a physical object being used to represent other physical objects.

6. Abstract Models. Consider a simple linear relationship between two variables, **x** and **y**, expressed by the equation

$$\textbf{(1)}\quad \mathbf{y = a\,x + b.}$$

This equation is a linguistic object, but also a physical object, letters on paper. But the relationship described is some sort of abstract object, more abstract than any written equation, which could use different letters or be written in another form, such as:

$$\textbf{(2)}\quad \mathbf{y - a\,x - b = 0.}$$

For the purposes of this exposition, I will take the existence of such abstract objects as unproblematic. We might call them *pure mathematical models* to distinguish them from what are more commonly called mathematical models, which I would then call *applied* mathematical models. More about these shortly.

This relationship can also be presented graphically, as in Figure 5. What are we to say about this graph? It is, I would say, a physical counterpart of the abstract model of the same linear relationship, that is, a physical model of a linear relationship. Like all physical models, of course, it is imperfect, and thus at best only similar to the abstract model.

[Insert Figure 5 here]

Beginning with the *pure* mathematical model, we can construct an *applied* mathematical model by replacing its mathematical elements with models of real objects and relations. For example, we can create a general model in which the variable **y** is distance from a fixed origin, **x** is time **t** from an arbitrary starting time, which can be zero, **a** is the velocity **v** of a moving point, and **b** the initial distance $\mathbf{d_0}$ of the moving point from the origin.

We can then create a still more specific model, say of an auto moving away in a straight line from an intersection at velocity **v** having started at time zero a distance $\mathbf{d_0}$ away. Here we are talking about **models** of an auto and of an intersection. In the model, the auto travels in a perfectly straight line at a perfectly constant velocity. Its distance from the idealized intersection at any time is then given by the equation

$$\text{(3)}\quad \mathbf{d(t) = v\,t + d_0.}$$

One may say that, in the model, this equation is *true*. What one cannot say is that the equation is true of the position of a real auto. No real auto can maintain a genuinely constant velocity in a perfectly straight line. The question, as always, is how similar the real situation is to the model of the situation.

Here one may complain that I am creating models beyond necessity. One must judge this complaint in light of the traditional way of handling the undeniable fact that no real objects exactly satisfy any simple mathematical relationships. The traditional way is to introduce *margins of error* into equation. Thus equation (3) becomes

$$\text{(4)}\quad \mathbf{d \pm \delta_d d = (v \pm \delta_v v)\,(t \pm \delta_t t) + (d_0 \pm \delta_d d_0),}$$

and this equation may indeed be true of the real auto. In this way one can preserve the idea that representation in science is to be understood solely in terms of the truth of statements.

While technically correct, this is not necessarily the best way of interpreting the actual use of abstract models in the sciences. The margins of error rarely appear in the descriptions or calculations until one gets to the point of comparing theoretical predictions with actual measurements. This practice strongly supports interpreting the original equations, without explicit margins of error, as referring not to actual things but to abstract models of which they are true by definition. When it comes time to compare the abstract model with reality, the deltas may then be understood as specifying the degree of similarity (either expected or actual) between the abstract model and the real system.

On this view, mathematical modeling is a matter of constructing an idealized, abstract model which may then be compared for its degree of similarity with a real system. The tendency to identify the model with the equations used to define it are then seen as a holdover from an excessively positivistic view of science that attempted to avoid abstract entities and identified underlying structures with their observable manifestations, such as minds with behavior, probability with relative frequency, and theories with their linguistic formulations.

7. Hypotheses. There is a use for statements like (4) above, although it is a little different from what is usually thought. Equation (4), as I understand it, is a *hypothesis* about a particular system in the real world. As such it may be judged true or false based on evidence obtained by examining the real system. In more general terms, statements like (4) say that some particular real system is similar to the proposed model, to the degree specified by the various error terms. It is also possible to generalize statements like (4) to cover classes of systems. In this case, the error terms might be explicitly specified, or they may be place-holders for the specific error terms in the various individual hypotheses included in the generalization.

8. Theoretical Models. In a slight departure from my earlier (Giere 1988) uses of the term *theoretical model*, I now want to reserve it for a special class of abstract models, those constructed with the use of what I shall call *theoretical principles*. Examples of such principles are Newton's Laws, The Schroedinger equation in quantum mechanics, the Principle of Relativity, the Principle of Natural Selection, and the laws of Mendelian Genetics.

There has long been a debate in the philosophy of science as to whether Newton's Laws, for example, should be understood, on the one hand, as definitions or conventions or, on the other hand, as empirical claims, either universal generalizations or claims of natural necessity. My view is that the initial question is defective because it presupposes that Newton's Laws are genuine statements which must be either true or false. The question is then one of the source of the truth or falsity. If one rejects this presupposition, one is free to answer, "both".

The issue may be posed as one of the status of Newton's term *body*. Does this term refer to empirical objects such as cannon balls and planets, or does it refer to abstract objects. I suspect Newton himself thought it referred to real physical objects. But we can make better sense of what he was doing, I claim, if we take him instead to be referring to abstract objects. These objects are then *defined* as things satisfying the three Laws of Motion plus the Law of Universal Gravitation. So, on this interpretation, Newton's Laws are definitions. One can then explore, as a mathematical exercise, the characteristics of various systems of bodies, such as two bodies moving in three dimensional space subject only to the three laws plus gravitation. Here I would say one is exploring features of a particular type of theoretical model. No empirical claims are being made, only claims about the model, which, if true, are true by definition.

However, if one identifies various real objects, such as the Earth and Moon, as bodies, then one has a model of this particular real system. In this case, one can formulate empirical claims as theoretical hypotheses about how the real system should behave if it is indeed similar to the model in the requisite respects. So the laws can be used to make empirical claims even though, taken by themselves, they imply no such claims.

Newton's gravitation law is usually rendered as "All bodies attract ...." In this form it sounds like an empirical generalization. And perhaps that is how Newton himself conceived of it. But a slight change to "Bodies attract ...." sounds like it could be part of a definition of bodies, on the order of "A circle is a plane figure ...." That is how I suggest we think of the law of universal gravitation. There are independent reasons for preferring this interpretation.

For one, Newton's bodies are said to be mass points. No real object can be a mass at a point. Any real thing with (classical) mass must be somewhat extended. To apply Newtonian models to real objects one must treat their mass as being concentrated at their "center of mass," which ideally is a point. This supports the interpretation of Newton's laws as *defining* idealized abstract objects rather than as *describing* real objects.

Second, there can be no direct evidence for the gravitation law understood as an empirical generalization. Even for as few as three bodies subject only to their mutual gravitation, the equations of motion admit of no exact solution. One must make simplifying assumptions, which, in my view, is to create simplified models known to be less than perfectly similar to a real three-body system. What we do have is a great number of one and two-body models that have been shown to be very similar to real systems. Textbooks in mechanics are devoted to developing a variety of such models and exploring their mathematical properties.

9. Mathematical Modeling. The practice of mathematical modeling in various areas illuminates the contrast between theoretical models and merely abstract models. Much mathematical modeling proceeds in the absence of general principles to be used in constructing models. Rather, one has a number of different *mathematical techniques* useful for constructing models, such as differential equations and systems of linear equations. These are deployed as the situation requires.

For example, in modeling the growth of organic populations, as in ecology, two sorts of models one may employ are exponential models and logistic models. The former go to infinity with time; the latter level off to a finite limit, as shown in Figure 6. The latter better represent most real populations over extended periods of time. Here there are no over-arching theoretical principles at work.

[Insert Figure 6 here]

If, however, one wishes to model changes in gene frequency in a population, then one employs principles of genetics, such as Mendel's laws, and not merely mathematical techniques, in the construction of well fitting models. One such model is characterized by the well-know Hardy-Weinberg Law.

Here, by the way, it is obvious that one has models and not literal descriptions. The mathematics of population growth and change typically employs continuous variables. But populations consist of discrete individuals, so their growth cannot literally be continuous. Nevertheless, for even just moderately large populations, continuous models may fit very well. In population genetics one finds people saying that they are assuming an *infinite* population. Here it is abundantly clear that they are talking about models and not about any real populations.

9. Models and Theories. In the introduction and several papers in a forthcoming collection of papers entitled *Models as Mediators*, Mary Morgan and Margaret Morrison argue that models should be though of as somewhat autonomous agents operating the region between data and theories. Much of what they say about the *uses* of models in science strikes me as quite illuminating. But they say very little about what models *are* and even less about what theories might be. They give the impression of presuming the traditional view of theories as sets of statements. They also seem to take for granted that models are typically quite inexact and oversimplified in known ways while theories are much less so. They even suggest that as models are refined, they may *become* theories. This seems to imply that models are also linguistic entities. How could something not a statement become a statement?

I think most of these latter ideas are mistaken. On my interpretation, the model/theory distinction is mainly a reflection of the extent to which a branch of inquiry is guided by broad general principles. Where there are such principles, as in many areas of physics and biology, the models employed often, though not always, embody these principles. Where such principles are lacking, the models employed derive principally from various mathematical techniques. In both cases, however, reasoning about the world is primarily reasoning with models. It is models almost all the way up.

10. Models and Data. Two years after publishing his paper on the meaning of models in empirical science, Suppes published an equally influential paper titled: Models of Data (1962). A central message of this paper was that higher level models are not compared directly with data, but with models of data which are lower down in a hierarchy of models. A similar point has been made in recent years by James Woodword (1989) who insists that is not data but phenomena that theories explain and that are used to test theories. And phenomena are first constructed from the data. Statistical techniques, for example, are among the primary means of constructing models of data from data.

On this view, when testing the fit of a model with the world, one does not compare that model with data but with *another model*, a model of the data. Thus one reasons from a high level model not to predictions about data, but to predictions about a model of possible data. The actual data are processed in various ways so as to fit into a model of the data. It is this latter model, and not the data itself, that is used to judge the similarity between the higher level model and the world. In between, as Suppes already insisted, must be a *model of the experiment*. A version of Suppes' hierarchy of models is shown in Figure 7. It is models almost all the way down.

[Insert Figure 7 here]

11. Models and Analogies. Finally, one can also provide an account of the role of analogy in science based on representational models rather than on instantial models. Scientists have at their disposal an inventory of various known phenomena and the sorts of models that fit these phenomena. When faced with a new phenomenon, scientists may look for known phenomena that are in various ways similar to, which is to say, analogous with, the new phenomenon. Once found, the sorts of models that successfully accounted for the known phenomena can be adapted to the new phenomenon. In the process, features of the old models may suggest unknown features of the new phenomenon. So reasoning by analogy can be a fruitful means to new discoveries. I would say, however, that such analogies are suggestive only, and provide very little reason actually to believe that the suggested features will in fact be found in the new phenomenon. That, I think, always requires independent confirmation.

12. Conclusion. By way of conclusion, let me present one final contrast between the instantial view of models and the representational view. On the instantial view of models, there are direct relationships of reference and truth between linguistic expressions and objects, as pictured in Figure 8. This conception works fine for well-defined abstract objects. It does not work so well for physical objects, whose natures and relationships are not so well-defined.

[Insert Figure 8 here]

On a representational conception of models, language connects not directly with the world, but rather with a model, whose characteristics may be precisely defined. The connection with the world is then by way of similarity between a model and designated parts of the world, as shown in Figure 9.

[Insert Figure 9 here]

In sum, scientific reasoning is to a large extent model-based reasoning. It is models almost all the way up and models almost all the way down.

References

Borges, J.L., 1954, *Historia Universal de la Infamia*, Emece, Buenos Aires.

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

Hodges, W., 1993, *Model Theory*, Cambridge University Press, Cambridge.

Morgan, M. and Morrison, M., forthcoming, *Models as Mediators*, Cambridge University Press, Cambridge.

Suppes, P., 1960, A comparison of the meaning and uses of models in mathematics and the empirical sciences, in: *Studies in the Methodology and Foundations of Science*, P. Suppes, ed., D. Reidel, Dordrecht, 1969.

Suppes, P.,1962, Models of data, in: *Logic, Methodology and the Philosophy of Science: Proceedings of the 1960 International Congress*, , E. Nagel, P. Suppes and A. Tarski, eds, 252-61, Stanford University Press, Stanford, CA.

Woodward, J., 1989, Data and phenomena. *Synthese* 79:393-472.