Re-presenting Scientific Representation

Roman Frigg

London School of Economics

Thesis submitted to the University of London
for the completion of the degree of a
Doctor of Philosophy

September 2003

Abstract

It is now part and parcel of the official philosophical wisdom that models are essential to the acquisition and organisation of scientific knowledge. It is also generally accepted that most models represent their target systems in one way or another. But what does it mean for a model to represent its target system? Surprisingly, this issue has hardly been recognised, much less seriously discussed. In the first part, I introduce the problem of scientific representation and argue for its importance. In the second part, I provide a critique of the current orthodoxy, the semantic view of theories. Though writers in this tradition do not explicitly address the issue of scientific representation, the semantic view implies that a model represents by being isomorphic or, in another version, similar to its target. I argue that this view faces insurmountable problems because both isomorphism and similarity are notions too weak to endow a model with representational power. In the third part, I develop a theory of representation that overcomes the shortcomings of the semantic view. The leading idea consists in taking representation to be explained in terms of three relations: *denotation*, *display* and *designation*. A model denotes its target system in roughly the same way in which a name denotes its bearer. At the same time it displays certain aspects, that is, it possesses these aspects and a user of the model thematises them. Finally, an aspect of the model designates an aspect of the target if the former stands for the latter and a specification of how exactly the two relate is provided.

For Mum and Dad

Contents

Part III – Re-Presenting Scientific Representation

Part I

Scientific Representation as a Philosophical Problem

Chapter 1

Models and the Three Conundrums of Scientific Representation

1. Models Matter – A Brief Survey

Models are of central importance in many scientific contexts. The roles the bag model of quark confinement, the billiard ball model of a gas, the Bohr model of the atom, the Gaussian-chain model of a polymer, the Lorenz model of the atmosphere, the Lotka-Volterra model of predator-prey interaction, the double helix model of DNA, or general equilibrium models of markets play in their respective domains are cases in point. Scientists spend a great deal of their time on building, testing, comparing and revising models and in many journal articles the bulk of the discussion is concerned with introducing and putting to use a particular model. In brief, models lie at the heart of modern science.

While their importance in actual research is beyond dispute, models had a rather fluctuating fate in the philosophical debate about science. For a long time it was a widespread conviction among philosophers that models were otiose in a systematic exposition of a scientific theory. Pierre Duhem took a science based on models to be inferior to one which presents its subject matter in the form of abstract and general principles (1906, Ch. 4) and the logical positivists regarded models as mere epiphenomena of scientific research. Rudolph Carnap famously remarked that 'the discovery of a model has no more than an aesthetic or didactic or at best heuristic value, but it is not at all essential for a successful application of the physical theory' (1938, 210) and Carl G. Hempel held that 'all reference to analogies or analogical models can be dispensed with in the systematic statement of scientific explanations' (1965, 440). Though some writers, in particular Richard Braithwaite (1953, Ch. 4) and Ernest Nagel (1961, Ch. 6) tried to canvass a more favourable picture of the use and function of models, the positivist take on the subject matter remained deflationary.

The tides changed in the sixties of the last century. The positivist picture of science construes theories as families of sentences. This outlook is now referred to as 'syntactic view' or 'received view' of theories. It came under attack with the rise of the so-called semantic view of theories, by which it was eventually overthrown. On this view, a scientific theory is a collection of models rather than sentences, where models are construed as non-linguistic entities. The details of this extended debate are discerningly surveyed in Fred Suppe (1977, 3-241) and need not concern us here. The salient point in connection with the replacement of the syntactic by the semantic view is that the latter, by construing theories as families of models, assigns models a place right at the heart of science. Over the years, the semantic view has been developed in different ways by, among others, Fred Suppe (1988), Ronald Giere (1988), Patrick Suppes (1960a, 1967, 1970) and Bas van Fraassen (1980, 1989, 1991), and has eventually become the currently dominant view on the nature of scientific theories.

Closely related to the semantic view is the German structuralist programme, which also adopts a model-based view of scientific theories. The programme has been developed over many years by a group of philosophers around Wolfgang Stegmüller in Munich and has found its classical formulation in a monumental treatise by Wolfgang Balzer, Ulises Moulines and Joseph Sneed (1987). The state of the art of the programme is documented in a special volume of *Synthese* (Moulines *et al.* 2002).

Parallel, but by and large unrelated to the rise of the semantic view, Peter Achinstein (1968, Chs. 7 and 8), Max Black (1960), Mary Hesse (1963) and later on Nancy Cartwritght (1983) have argued in one way or another for the importance of models on the grounds that they are indispensable to scientific practice, in its experimental as well as theoretical aspects. If models play an essential role in actual research, we cannot dismiss them as irrelevant in a philosophical account of what science does. Finally, about two decades on, a group of philosophers gathered in a spirit similar to this under the umbrella of the 'models as mediators' project and embarked on a detailed study of how models function in different investigative contexts (see the essays collected in Morgan and Morrison 1999).

This enumeration is not meant to suggest that all the aforementioned philosophers employ the same notion of a model, let alone that they share a common philosophical programme. In fact, there is little consensus as to what models are, how they function, and how they relate to both theory and reality. However, despite these differences, there is one point that all share in common: models matter. What the above survey is supposed to show is that nowadays philosophers across the board assign models a central place in science. Whatever picture of what science does, what its aims are and what it achieves is canvassed, models are an integral part of it. Much remains to be said about the differences between diverging accounts, but for now it is the common denominator on which I would like to lay stress and which I take as my point of departure: we can't do without models.

2. Description versus Representation

Fashion changes, in philosophy as well as on the catwalk. So is the emphasis on models more than a quirk? I think so. The move from the received view to a model-based account of science marks a fundamental change, namely the transition from a descriptive to representational understanding of science: science does not describe, it represents its subject matter.

On the received view, science aims at providing theories, and these are taken to be families of sentences. Thus construed, scientific theorising is a linguistic activity aiming at giving us a description of whatever we investigate. Consequently, understanding science amounts to understanding how the language of science works. So we find among the hotly debated problems at the time issues like the nature of theoretical terms and the character of correspondence rules. Although there is little agreement on actual answers, the 'linguistic paradigm' itself enjoyed univocal acceptance for a long time: understanding science is understanding how its language works.

This changed radically when models entered the scene. Though views on what exactly they are vary widely, it is generally held that models are something other than a set of sentences, i.e. that models are non-linguistic entities. This 'anti-linguistic turn' in the philosophy of science originated with the pioneers of the semantic view and found its programmatic expression in van Fraassen's remark that the main lesson of twentieth century philosophy of science may well be that no essentially language-dependent concept has any philosophical importance at all (1980, 56). One may not want to go as far as that, but the consequence is inevitable: if the main units of scientific theorising are non-linguistic, then description has to be replaced by something else, which I take to be representation. Words describe and models represent.[1] Hence, the shift from a sentence to a model-based account of science amounts to replacing a descriptive view of science with a representational one: science gets a grip on selected parts of the world not by describing, but by representing them in a model.[2]

Many philosophers, realists and antirealists alike, agree with a characterisation of science as an activity aiming at representing the world, among them Cartwright (1999, esp. Ch. 8), Giere (1988; 1999; 2002), Hughes (1997), Kitcher (1993), Morgan and Morrison (1999), Morgan (1999a), Morrison (1999), Psillos (2000), Redhead (2001), Suppe (1989), Suppes (1967; 1970), van Fraassen (1980; 1997; 2002), to mention just a few. Needless to say, views on what representing the world amounts to vary widely; I come to that later. But before moving on I should guard against a common misconception of representation. According to a popular myth, to represent an item amounts to giving something like a mirror image of that item; that is, a representation is thought of as a copy or imitation of the thing represented. In short, representation is wedded to realism. This is a mistake. Representations can be

[1] The term 'representation' can be used in a broad and in a narrow sense. In the broad sense it refers to anything that stands for something else and therefore encompasses, among many other things, also language. In a narrow sense it covers only cases of non-linguistic representations such as paintings, photographs, maps, diagrams or, last but not least, scientific models (I do not want to suggest that the notion of representation involved in all these cases is the same). I use the term in the narrow sense unless indicated otherwise; in the latter case I make the broad meaning explicit by adding a qualification and talking, for instance, about 'linguistic representation'. When using 'representation' in this broad sense, one can describe the move from a sentence to a model based view as the recognition that science employs object-to-object rather than word-to-object representations.

[2] I do not hereby claim that *all* models are representational; some are merely tools for manipulating or intervening in the world. But many models do represent and that is my point of departure.

realistic, but they need not. There is nothing in the concept of a representation *per se* that ties it to realism. Human beings don't look like Alberto Giacometti's sculptures, yet they represent Humans; or the harmonic oscillator is not a mirror image of an actual pendulum bob, yet it represents this bob. Representation and realism are just two different issues. One can reject realism without rejecting representation. Representations, in science as well as in art, can be of different sorts; representation is a much broader notion than mirroring or depiction. Some representations are realistic, others are not. It would be a wrong move to deny the status of a representation to one or the other of these. One of the things we expect from a theory of representation is that it can account for the difference between realistic and non-realistic representations, a possibility we forgo if we dismiss one of them out of hand.

That representation and realism are distinct is of great importance to the project of this thesis. A theory of scientific representation is not, as one might mistakenly believe, a defence of scientific realism in disguise. On the contrary, in order to achieve an adequate understanding of scientific representation, it is necessary that the concept be freed from the stranglehold of realism. Representation as such presupposes neither realism nor antirealism.[3]

This said, I am now in a position to introduce the three main conundrums a theory of scientific representation has to deal with, the enigma of representation, the ontological puzzle and the problem of *quomodity*. But before that, a methodological remark seems in place. Throughout this thesis, I make extended use of examples from art, in particular sculpture and painting, and use ideas originating in philosophical aesthetics. This is purely heuristic and nothing in the final formulation of any of my views on scientific representation hinges on it. Those who are adverse to drawing on parallels between art and science can skip the respective passages and replace the artistic examples by scientific ones without missing anything essential to my view on scientific representation. Although it is important that one *can* do that (a theory of scientific representation has to be an independent theoretical unit!), I hope that no one *will*. The problems in the two fields are similar in many respects, which

[3] For this reason I use the term 'representation' rather than 'depiction', or 'imitation'. Despite some realistic overtones, 'representation', unlike its alleged synonyms, is neutral enough to accommodate realistic and non-realistic varieties of aboutness.

makes pointing out both analogies and disanalogies between them extremely productive. Some ideas from a theory of pictorial representation do *mutatis mutandis* carry over to the case of models, while others do not. In the former case we gain valuable insights we can build on, in the latter it is instructive see why they do not. The heuristic value of this parallel is further reinforced by the fact that while there is a rich and intricate literature on artistic representation, the problem of scientific representation has, by and large, been forgotten in recent debates – I come to this below. Why then not learn from philosophy of art what can be learned?

3. The Three Conundrums of Scientific Representation

Models are representations of something else, usually a selected part or aspect of the world (henceforth referred to as the 'target system'). But in virtue of what is a model a representation of something else? To appreciate the thrust of the question, consider the analogue problem in the visual arts. When seeing Camille Pissarro's *Boulevard des Italiens* we immediately realise that it depicts one of the glamorous streets of *fin de siècle* Paris. Why is that? The symbolist painter Maurice Denis famously took wicked pleasure in reminding his fellow-artists that a painting, before being a battle horse, a nude, or some anecdote, essentially is a plane surface covered with paint. A painting *as* a painting is a welter of lines and dots, a bounded collection of curves, shapes, and colours. Following Beardsley (1981, 267), I call a painting thus construed a 'visual design'. The puzzle then is this: due to what is a visual design a picture; how do lines and dots represent something outside the picture frame? Per se, *Boulevard des Italiens* is a canvass covered with a myriad of oil-paint brush strokes, yet it depicts a scene of urban life. How can a configuration of flat marks on a canvass do this? What is it for that canvass to be a pictorial representation of a Parisian street? This is what Flint Schier called the enigma of depiction (1986, 1).[4]

Like pictures, models are representations of something beyond themselves. This naturally raises the same question in a different setting: in virtue of what is a model a

[4] Different formulation of this problem can be found in Beardsley (1981, 266-68), Lyas (1997, 43), Peetz (1987, 227), and Schier (1986, 1-2).

representation of something else; what does its 'aboutness', its 'representational power', consist in? Or to render the question more precisely: what fills the blank in: '*M* is a scientific representation of *T* iff ___', where '*M*' stand for 'model' and '*T*' for 'target system'. Slightly altering Schier's congenial phrase, I refer to this as the *'enigma of representation'*. This is the first of three major conundrums a theory of scientific representation has to come to terms with. Presenting a solution to this conundrum amounts to specifying the general features that *every* scientific representation has to possess in order to be a scientific representation.

The second conundrum becomes palpable once we try to bypass the analogy with painting and give an explicit statement of the enigma of representation. The problem is that it is not clear what we are to substitute for phrases like 'a welter of lines and dots'. Specifying the 'material substratum' of a piece of art is an unproblematic matter. Neither do the specifications 'oil on canvass' or 'bronze cast' on the label next to an exhibit cause any bewilderment in the spectator, nor will the need to provide such information give curators sleepless nights. Either the matter is just obvious, or the conservation department will come up with an answer. Not so with scientific models. It is not clear what models are and views on that matter vary widely. Some take models to be structures in the sense of set theory, others regard them as concrete objects and yet others as idealised or abstract entities – and this by no means exhausts the possibilities. So what then are models? What is the 'scientific analogue' of the visual design? This is the second question a theory of scientific representation has to answer. I refer to it as the *'ontological puzzle'*.

The third conundrum is what I call the 'problem of *quomodity*', which comes in a factual and in a normative version. Not all representations are of the same kind; there are different ways in which models can represent reality. In the visual arts this is so obvious that it hardly deserves mention. Innumerable pictures have been painted portraying the same subject, a Provence landscape, say, each using different means and devices. There are ink drawings depicting the landscape by using thin and clearly marked lines to reproduce selected aspects of the relief's shape as they appear to the human eye from a certain point of view under certain circumstances. A pointillist painter works under the assumption that the human visual field can be decomposed into a myriad of coloured spots and achieves representation by reproducing – at least to some extent – on canvass the pattern of colours the object produces in the visual field. Yet others focus on simple geometric figures such as cubes and cylinders and represent their subject by first 'seeing' the subject as composed of such elementary figures and then reproducing these figures on canvass in both colour and form. Needless to say, this list can be extended at will.

This pluralism is not a prerogative of the fine arts. The representations used in the sciences are not all of one kind either, not even when they represent the same target system. Weitzäcker's liquid drop model represents the nucleus of an atom in a manner very different from the shell model. The latter represents the nucleus by means of part-to-part correspondence; it posits a set of entities that are matched with the constituents of the nucleus (the nucleons) and the claims we derive from the model are based on this part-to-part correspondence. Nothing of that sort happens in the liquid drop model. The model has no parts that could correspond to parts of the target. It represents some selected features of the nucleus as a whole by portraying its shape. Similarly, a scale model of the wing of an air plane represents the wing in a way that is different from how a mathematical model of its cross section does. Or Bill Phillips' famous hydraulic machine and Hicks' mathematical models both represent a Keynesian economy but they use very different devices to do so. Scientific representation is not a monolithic concept. We should not be fooled into thinking that all scientific representations work in the same way. On the contrary, as in painting, there seems to be a variety of devices one can use to achieve representation. But unlike in painting, it is not at all obvious what these devices are. This leaves us with the question of how scientific models represent. What devices are there and how exactly are they put to use? For want of a better term, I call the way in which a model represents the target system its '*quomodity*', from Latin *quomodo* = how (occasionally I use 'mode of representation' or 'representational strategy' as synonyms for '*quomodity*'). A theory of representation has to come up with a taxonomy of different *quomodities* and provide us with a characterisation of each of them. This is the factual aspect of the problem of *quomodity*.[5]

[5] The question 'how does a model represent its target?' seems to be ambiguous. One may say that on an alternative reading, 'as a system of pipes' or 'as a collection of billiard balls' are also answers to the question of how the Phillips machine or the Maxwell model represent an economy or a gas. But this way of answering the how-question is only seemingly different from specifying the representational devices a model employs. Answers like 'as a collection of billiard balls' are ellipses

A further aspect of the problem of *quomodity* is the normative question of whether we can draw a distinction between scientifically acceptable and unacceptable modes of representation. Do the epistemic interests of scientific enquiry constrain the choice of modes of representation? Some may be willing to grant that there are different representational strategies but still hold that only some of them truly deserves the label 'scientific'. A theory of representation should address the question of whether there are some *quomodities* that can be singled out as acceptable while others are have to be dismissed as unscientific.

In sum, a theory of representation has to come to terms with three conundrums, two semantic, and one ontological. The two semantic issues are the enigma of representation (*why* does a model represent its target?) and the problem of *quomodity* (*how* does a model represent its target?); the third issue is the ontological puzzle (*what* is a model?).

To frame the problem in this way is not to say that these are separate and unrelated issues, which can be dealt with one after the other in roughly the same way in which we first buy a ticket, walk to the platform and then take a train. This is not the case. This division is analytical, not factual. It serves to structure the discussion

---

and when spelled out they amount to the aforementioned specifications. Consider the gas example. Saying that the model is a collection of billiard balls is nothing but a convenient way of defining what the properties of the model are. Maxwell did not care about billiard balls *per se*, what matters is that we deal with entities which have certain dynamical properties, interact with hard walls in a certain way, and so on. Since we naturally associate all these properties with billiard balls, saying that the model consists of billiard balls is an efficient way of specifying the properties of the model. Over and above this, billiard balls don't do any work and the locution 'as a set of billiard balls' can be replaced by an enumeration of the mechanical properties of billiard balls without any loss. Then, the claim that the billiard balls *represent* the gas amounts to saying that the properties the balls possess represent the properties of the gas in a certain way. But this implies an answer to the question about representational devices. Specifying how certain properties of the model represent certain properties of the gas amounts to knowing how these two sets of properties relate to one another and this in turn amounts to knowing what representational strategy is used. This, again, becomes clear from the analogy with painting. We understand how a painting represents once we know what representational strategy (e.g. sameness of colour) is used to relate the features of the painting (e.g. the colour patterns it exhibits) to features of the object depicted. In the same vein we understand how a model represents once we know what strategy is employed to relate the properties of the model to the properties of the target.

and to assess suggestions but it does not imply that an answer to one of these questions can be dissociated from what stance we take on the other issues. Let me illustrate this by dint of the illusion theory of pictorial representation. On this view, the enigma of representation is answered by an appeal to illusion: X pictorially represents Y iff X causes the illusion of seeing Y in the spectator. A picture represents Piccadilly Circus if it makes us believe that we see Piccadilly Circus. But how does it achieve this? This is the problem of *quomodity*. What devices does the picture employ to bring about the illusion? Is it the use of lines to portray visible shapes, is it sameness of colour, both of them together, or yet something else? And this answer cannot be given without presupposing a certain ontology. If the picture represents by dint of sameness of colour and shape, say, it must be true, trivially, that the picture belongs to the kind of objects that can have colours and display shapes.

4. Taking Stock

Where does the debate on scientific representation stand? The answer to this question is rather sobering. In many respects the debate has hardly started. Despite its importance, the issue of scientific representation has barely been recognised, much less seriously discussed. The question of how an *object* represents another one has, by and large, gone unnoticed in twentieth century analytical philosophy of science.[6] Unlike word-to-object representation, object-to-object representation still awaits detailed treatment informed by the practices of representation in the sciences.

The different problems within a theory of representation have met with different degrees of inattention. Among the three basic conundrums, the enigma of representation has the poorest track record. Within analytical philosophy of science, it has not even been recognised as a problem and virtually nothing has been written on it. Given that models have been with us for at least forty years now, this is rather surprising because how one thing comes to represent another one seems an important

[6] Two noteworthy exceptions are Hughes (1997) and Suárez (1999, 2002). I discuss their views in Chapter 7.

and puzzling question. Accordingly, a fair portion of this thesis is dedicated to this issue.

The ontological puzzle, though as marginal a topic as the enigma of representation, has received some tangential mention in parts of the literature on the semantic view of theories. Many authors writing in this tradition are strongly influenced by formal approaches to science and in this vein take models to be structures in the sense of set theory (references will be given in Chapter 2). Although this view is not explicitly put forward as an answer to the question about the ontology of models, it does not seem farfetched to take it as such.

Unlike the other two conundrums, various discussions have been going on for a while which quite naturally can be understood as at least partially addressing the problem of *quomodity*. What I have in mind are the extended debates about the nature of idealisation, the functioning of analogies, and the like. Though these issues have not been discussed within the context of a theory of scientific representation, it seems quite natural to understand idealisation, for instance, as an answer to the question of the *quomodity* of a model. The problem with the issue of *quomodity* is a lack of systematisation rather than a lack of attention. Icons, idealised models and analogies are not normally discussed under one theoretical umbrella. As a consequence, we lack comparative categories that could tell us what features they share and in what respects they differ. What we are in need of is a systematic inquiry, which provides us with both a characterisation of individual *quomodities* as well as a comparison between them. What, for instance, makes an icon an icon, and how does it differ from an analogy? In tackling this problem we can build on intuitions, but they need to be systematised and transformed into a coherent and comprehensive web of notions characterising how models face reality.

The bottom line is this: the enigma of depiction and the ontological puzzle need to be recognised and discussed; existing accounts of *quomodity* need to be systematised and developed.

5. Why Bother?

I have presented the lack of a philosophical literature on scientific representation as a lacuna that calls for action. But a critic might counter that this, rather than being a call for action, is indicative of the irrelevancy of the topic. Needless to say, I disagree. If it is worthwhile asking – and I think it is – how words acquire meaning and reference, how speech acts work, or what it means for a sentence to be true, I cannot see why it should not be worthwhile asking how models represent. Representation it is no less a philosophical puzzle than reference or truth and as such deserves our attention. But this, I anticipate, may not satisfy the critic. For this reason I will now sketch how representation relates to the epistemology of science, in particular to issues such as the acquisition and the nature of knowledge and the character of scientific explanation.

Science aims at giving us empirical knowledge. If models are to serve this purpose, they must be representational. There is an intimate connection between knowing and representing. Cognitive acquisition in science is often mediated by representation. This interdependence is grounded in the fact that models are the units on which significant parts of scientific investigation are carried out rather than on reality itself. We study a model and thereby discover features of the thing it stands for. For instance, we study the nature of the hydrogen atom, the dynamics of populations, or the behaviour of polymers by studying their respective models. But if models are to give us knowledge about the world, they must be representational. They can instruct us about the nature of reality only if we assume that (at least some of) the model's aspects have counterparts in the world. Hence, in order to be a source of knowledge models must be representational.

What kinds of knowledge do models provide us with? If the acquisition of knowledge is closely tied to representation and if there are different ways of representing, then there are different ways of knowing. We learn from a model by first inquiring into the features of the model itself and then transferring the findings to the real-world system it represents; that is, we 'export' the knowledge we acquire about the model to its corresponding system in the world. But in order to know how and in what way knowledge about a model can be converted into knowledge about a system, we have to know what sort of representational relation holds between the two. If, for instance, we have a model we take to be a realistic depiction, this transfer is accomplished in a different manner than when we deal with an analogue, or one that involves idealising assumptions. Thus the way in which a model represents its target directly bears on what we can learn from it about physical reality – different ways of representing, different ways of knowing.

What are these ways? To answer this question we need to know what kinds of representations there are. That is, we need to have a solution for the problem of *quomodity*. Only when we are clear on different representational strategies can we start discussing the question of what kinds of knowledge they provide us with.

Finally, there is a close connection between knowing and understanding, as well as knowing and explaining. Different kinds of knowledge correlate with different notions of understanding and with different types of explanation. Therefore, the way in which a model leads to understanding and the kinds of explanations it provides us with is a function of how it represents. For this reason, the issue of representation has a direct bearing on explanation and understanding.

6. Aims and Plan

The aim of this thesis is to provide comprehensive answers to both the enigma of representation and the ontological puzzle, along with an indication of the direction in which an answer to the problem of *quomodity* has to be sought. Within the limits of the current PhD thesis it is not possible to present a complete answer to the problem of *quomodity* as well. Such an answer, I maintain, has to be taxonomic. That is, what we have to come up with in response to the problem of *quomodity* is a catalogue of the different ways in which models can represent their targets, along with an account of how these ways differ from one another. As far as I can see, there is no 'a priori' method of compiling such a catalogue; it is extended case studies that will have to tell us in what way different models face reality. I discuss one particular case in Chapter 7 in order to indicate what the study of representational strategies might look like; but to come up with a (more or less) comprehensive set of cases is beyond the scope of this thesis.

This gives rise to the following plan. The thesis has three parts. Part I, the present chapter, introduces the problem of scientific representation and argues why it is important. Part II is a critique of the current orthodoxy. Although the issue of representation has received little explicit attention within philosophy of science so far, the currently dominant view on the nature of scientific theories, the so-called semantic view, deals with issues that are relevant to a theory of representation. So I begin by extracting a consistent account of representation from its writings, which I then criticise as untenable. In Part III, finally, I develop a positive view of how scientific representation works, which overcomes the shortcomings of current accounts.

Part II

Why Current Accounts are Blind Alleys

Chapter 2

Strictures on Structures

1. The Structuralist Conception of Models

One influential account of models is articulated within the semantic view of theories, which is now, in one way or another, held by many philosophers of science. The problem in discussing this account is that the semantic view does not address the issue of representation, although it touches upon topics that are relevant to it. The explicit involvement with the problem of representation remains confined to a casual mention of the term and some brief remarks to the side every now and then.[7] So my first task is to extract a consistent account of representation from the writings on the semantic view.

At the core of the semantic view lies the notion that models are structures (where a structure, roughly speaking, is a collection of objects along with the relations in which they enter) and that they represent due to their being isomorphic to their target systems. For this reason, I refer to this as the *structuralist view of models.*[8] Another version of the semantic view has it that similarity rather than isomorphism is the correct specification of the relation obtaining between model and target. I focus on the version that deals with isomorphism in this chapter and the following, and defer a discussion of the similarity variant of the view to Chapter 4.

[7] Recently, Bas van Fraassen (2002b) and Steven French (2002) have paid some attention to the issue of representation within the framework of the semantic view of theories. However, no systematic account of representation emerges from their discussions.

[8] This coincides with the terminology used by its advocates. While the term 'structuralism' has been used by the German School all along, anglophone philosophers of science in the past preferred to refer to their position as the 'semantic view of theories'. More recently, however, some of them – van Fraassen, DaCosta, French, Ladyman, and Bueno – began to refer to their positions as 'structuralism'.

A structure (sometimes 'mathematical structure' or 'set-theoretic structure') S is a composite entity consisting of the following ingredients: (i) a non-empty set U of individuals called the domain (or universe) of the structure S, (ii) an indexed set O (i.e. an ordered list) of operations on U (which may be empty), and (iii) a non-empty indexed set R of relations on U.[9] Often it is convenient to write these as an ordered triple: $S=\langle U, O, R\rangle$.[10, 11] Note that nothing about what the objects are matters for the definition of a structure – they may be whatever one likes them to be. Similarly, operations and functions are specified purely extensionally; that is, n-place relations are defined as classes of n-tuples, and functions taking n arguments are defined as classes of $n+1$-tuples (I come back to this point in greater detail in the next chapter).

The crucial move now is to claim that scientific models are nothing but structures in this sense. In this vein Suppes declares that 'the meaning of the concept of model is the same in mathematics and the empirical sciences.' (1960a, 12) and van Fraassen posits that '[a]ccording to the semantic approach, to present a scientific theory is [...] to present a family of models – that is, mathematical structures offered for the representation of the theory's subject matter' (1997, 522). In short, models are structures.[12] This has become the cornerstone of the semantic view of theories.[13]

---

[9] This definition of structure is widely accepted among mathematicians, logicians, and formally minded philosophers of science. See for instance Machover (1996, 149), Bell and Machover (1977, 9, 49, 162), Hodges (1997, 2), Boolos and Jeffrey (1989, 98-9), Solomon (1990, 168), Rickart (1995, 17), Shapiro (2000, 259), Bourbaki (1957, 12).

[10] Structures of this kind are sometimes also referred to as 'first order structures' because they are used to provide first order languages with a semantics. However, their use is not confined to first order logic. As Shapiro (1991, Ch. 3) points out, also second order languages can be interpreted in terms of the same structures. For this reason I refer to S as simply as a 'structure' and drop the qualification 'first order'.

[11] In what follows I will, for the sake of simplicity, not deal with operations. This is justified because (a) the set of operations can be empty by assumption and (b) ultimately operations reduce to relations (see Boolos and Jeffrey 98-99; Shapiro 1991, 63).

[12] Further explicit statements of this view include: Da Costa and French (1990, 249), Suppes (1960b, 24; 1970, Ch.2 pp. 6, 9, 13, 29), and van Fraassen (1980, 43, 64; 1991, 483; 1995, 6; 1997, 528-9; 2001, 32-3). This is not to deny that there are differences between different versions of the semantic view. The precise formulation of what these models are varies slightly from author to author. A survey of the different positions can be found in Suppe (1989, 3-37). How these accounts differ from one another is an interesting issue, but for the present purposes nothing hinges on it. As Da Costa and

However, in themselves structures do not represent anything in the world. They are pieces of pure mathematics, devoid of empirical content. But a representation must posses 'semantic content', that is, it must stand for something else. Structuralists can reply to this charge that structures, though not representations in themselves, are endowed with representational power when the relation they bare to their intended target system is specified. What is the nature of this relation? In keeping faithful to the spirit of the semantic view, the most natural choice is structural isomorphism. We then obtain: a structure S represents a target system T iff they are structurally isomorphic.[14]

This is in need of qualification. Assume that the target system T exhibits the structure S_T=<U_T, O_T, R_T>. Then consider the structure S. An isomorphism is a mapping f: $U_T \rightarrow U$ such that (i) f is one-to-one (bijective), (ii) f preserves the system of relations in the following sense: the elements $a_1, \ldots, a_n$ of S_T satisfy the relation R^T iff the corresponding elements $b_1=f(a_1), \ldots, b_n=f(a_n)$ in S satisfy R, where R is the relation in S corresponding to R^T. And similarly, for all operations g^T of S_T we have $f[g^T(a_1, \ldots, a_n)]= g(f(a_1), \ldots, f(a_n))$ where g is the operation in S corresponding to g^T. Note that the notion of isomorphism as introduced here is symmetrical, reflexive, and transitive: if A is isomorphic to B, then B is isomorphic to A; every structure is isomorphic to itself; and if A is isomorphic to B and B to C, then A is isomorphic to C.

French (2000, 119) – correctly, I think – remark, '[i]t is important to recall that at the heart of this approach [i.e. the semantic approach as advocated by van Fraassen, Giere, Hughes, Lloyd, Thompson, and Suppe] lies the fundamental point that theories [construed as families of models] are to be regarded as *structures*.' (original emphasis)

[13] As introduced here, models *are* structures, no more and no less. I should mention, however, that this use of the term 'model', although motivated by mathematical logic, differs from how it is used there. The difference is that in the case of mathematical logic linguistic elements are considered part of the model as well. More specifically, a model is commonly taken to consist of three elements: first, a structure defined as a set of objects endowed with relations and operations, second, a 'basic' language consisting of various symbols and, third, a denotation-function which assigns these symbols appropriate elements (or sets thereof) in the structure (see for instance Machover 1996, Ch. 8; Hodges 1997, Ch. 1).

[14] This view is extrapolated from van Fraassen (1980, Ch. 2; 1989, Ch. 9; 1997), French and Ladyman (1999), French and Da Costa (1990), French (2000), and Bueno (1997 and 1999), among others. Van Fraassen, however, adds pragmatic requirements – I shall come to these below.

In sum, the structuralist view of models is the following: a model *M* is a structure; and *M* represents a target system *T* iff *T* is structurally isomorphic to *M*.[15]

Mauricio Suárez (1999, 79) has raised the objection against this view of models that the capacity of a model to represent must be an inherent part of it and not something that is added to it as an 'external factor'. More specifically, since structures *per se* do not represent anything in the world, it is inappropriate to refer to them as models. Models inherently 'point' to their targets and do not need to be connected to them by postulating an external relation (structural isomorphism) to hold: '[...] a model is a representation, as it essentially intended for some phenomenon; its intended use is not an external relation that we can choose to add to the model, but an essential part of the model itself.'

I agree with this. Taking a model to be a structure, and nothing but a structure, is like taking a painting to be merely a collection of brush strokes on canvas. However, proponents of the semantic view could now counter that this argument rests on a confusion about the use of the term 'model'. They may say that when they speak of models as structures what they actually mean is 'structure *plus* isomorphism'. On this reading, the structuralist view is that a model is a structure *plus* the isomorphism that holds between the target and the structure – and if at times bare structures are called 'models' this is just sloppy talk. We then obtain the following as a definition of the structuralist view of models:

(SM) The structure *S* represents the target system *T* iff *T* is structurally isomorphic to *S*; given that this is the case, the model *M* is the pair consisting of the structure *S* and the isomorphism that holds between *S* and *T*.

This is illustrated in Figure 1.

[15] I should mention that although this notion of a model is usually employed in the context of the semantic view of theories, it is not necessarily bound to it. As Da Costa and French (2000, 120-1) rightly point out, an account that emphasises the independence of models from theory (such as Morrison's or Cartwright's) could, in principle, still adopt a structuralist view of what models are and of how they relate to reality. Hence, nothing in what follows hinges on any particular feature of the semantic view of theories.

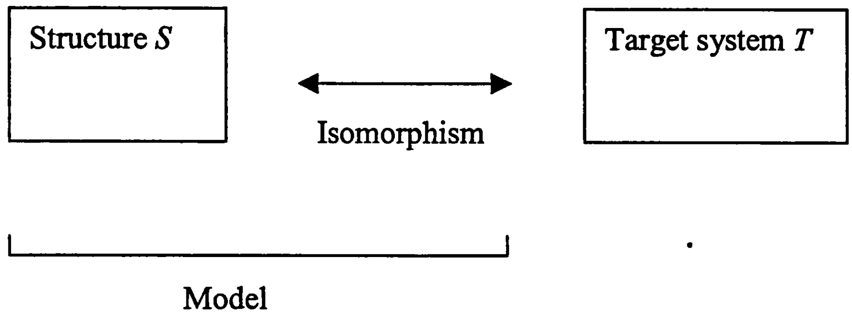

Fig. 1: The structuralist conception of representation.

The structuralist view of models comes in grades of refinement or sophistication. What I have presented so far is its simplest form. The leading idea behind all ramifications is to weaken the requirement that model and target have to be isomorphic. This can be achieved by replacing the isomorphism in the above definition by an embedding, a partial isomorphism, or a homomorphism, to mention just some options. However, I believe that the structuralist view is at its best at its simplest, and that refinements only obscure what it has to offer without reducing any of the serious difficulties that attach to it. For this reason, I consider the structuralist view in its simplest form throughout this chapter and confine my discussion of the ramifications to a brief section at the end, explaining why they do not fare better.

The aim of this and the next chapter is to argue that the structuralist conception of models is a blind alley. The arguments against it fall into two groups. Criticisms belonging to the first group, which I will be the dealing with in the present chapter, grant that the notion of an isomorphism between a model and its target is unproblematic, and proceed to showing that scientific representation cannot be explained in terms of isomorphism. Arguments belonging to the second group question this assumption. Isomorphism is a relation that holds between two structures and not between a structure and a piece of the real world *per se*. Hence, if we are to make sense of the claim that the model is isomorphic to its target we have to assume that the target exhibits a certain structure $S_T=\langle U_T, O_T, R_T\rangle$. This assumption is by no means trivial. What does it mean for a target system – a part of the physical world – to possess a structure? I discuss this question at length in Chapter 3 and argue that structural claims do not 'stand on their own' in that their truth rests on the truth of a more concrete description of the target system. As a consequence, concrete descriptions cannot be omitted from an analysis of scientific representation. For this reason one has to recognise that structures and isomorphisms alone cannot be a complete analysis of scientific representation.

Let us then, for the time being, assume that there is no problem with there being an isomorphism between a model and a target system. Given this, is (SM) correct? Before I can answer this question, I have to say what I mean by 'correct'. I take it that (SM) is correct if it provides us with a satisfactory response to the three conundrums of representation as introduced in the last chapter. What stance then does (SM) take in these matters? While things are clear as far as the ontological puzzle is concerned – models are structures – (SM) is ambiguous about the semantic issues. Is isomorphism supposed to be an answer to the enigma of depiction or to the problem of *quomodity* or to both? Since writers on the semantic view do not distinguish between these two problems, one can only speculate what their answer might be. In what follows I will discuss both possibilities and conclude that neither of them will do. As a response to the enigma of representation, isomorphism is patently inadequate, and it fares only slightly better when understood as an answer to the problem of *quomodity*. While isomorphism can be an answer to how some models represent, it is not the case that all models represent in this way, nor is it true that the presence of an isomorphism marks the difference between scientific and non-scientific forms of representation. Scientific representations can be isomorphic to their targets, but they do not have to be. So my conclusion will be that isomorphism does not provide us with an answer to either of the semantic issues. And what about the ontological claim? Are models structures? It will come as little surprise by now that I take this to be mistaken too, but I shall argue for this only in the next chapter.

I should end this introduction with a disclaimer. The criticism I put forward in this chapter and the next are only directed against structuralism as an account of scientific representation. I do not thereby claim that structures are unimportant or that structuralism cannot be the right route to go in other contexts. In fact, in Chapter 8 I argue for a structuralist answer to the problem of the applicability of mathematics. The point I am making in what follows is only that structures by themselves are not enough to make up a representational model.

2. Structuralism and the Enigma of Representation

Can (SM) account for why models represent? In this section I argue that it can't. To get a first idea of why it cannot, consider the following variation of a well-known thought experiment due to Hilary Putnam (1981, 1ff.).[16] Imagine a mathematician who is totally ignorant about physics. One day, sitting at the desk in his study, he writes down the equation $d^2x(t)/dt^2 = -k^*x(t)$ and discusses its solutions with all sophistication and rigour of a modern mathematician. A person with some background in physics immediately realises that this is the equation describing the motion of an oscillating pendulum bob; or using the model-theoretic jargon: the phase space trajectory the equation describes is a *model* of the motion of the pendulum bob. But did the *mathematician* (who, to repeat, is totally ignorant about physics) write down a model of the moving pendulum?

On a little reflection most of us would probably say: no, *he* didn't. What the *mathematician* wrote down is a piece of pure mathematics, a mathematical structure if you like, but not a model of any physical system. The mathematician, after all, does not know pendulums and hence could not have had the intention to come up with a model of one. He simply wrote down an equation and solved it. The fact that *someone else* can 'see' this structure as a model of a pendulum is not his concern.

The point I want to get at is that structural isomorphism is not sufficient to make something represent something else. In Putnam's original thought experiment an ant is crawling on the beach leaving a trace that resembles Winston Churchill. Does this trace represent Churchill? Putnam concludes that it does not, because '[s]imilarity [...] to the features of Winston Churchill is not sufficient to make something represent or refer to Churchill. Nor is it necessary.' (1981, 1)

Mutatis mutandis things are the same in the case of the ignorant mathematician. Unless he knows about oscillations and intends to come up with a model for them, his structures do not represent oscillations – regardless of whether they are isomorphic to oscillations or not.

---

[16] Black (1970, 104) discusses a very similar thought experiment.

Why not? I have used Putnam's story to draw attention to the problem of representation. But a story is not an argument. So the burden of proof seems to be on my side: I have to provide convincing arguments why isomorphism is not enough to establish representation. This is what I want to do in the remainder of this section, where I present four arguments for the conclusion that this view is wrong.[17]

But before I start arguing for my claim, I would like to use an observation from the history of science to further dispel the appearance of plausibility (or even obviousness) of an isomorphism-based view of representation. Many mathematical structures have been discovered and discussed long before they have been used in science. Hilbert spaces or continuous manifolds with non-zero curvature as studied in curvilinear geometry are cases in point. If we subscribe to the view that mathematics refers 'by itself' we have to accept that Riemann discovered general relativity or that that Hilbert invented the state space of a quantum system. This is obviously mistaken. It is only later that scientists started to *use* these mathematical tools to represent things and processes in nature.

(a) Isomorphism has the wrong formal properties

The first and somewhat simple reason why representation cannot be cashed out in terms of structural isomorphism is that the latter has the wrong formal properties: isomorphism is symmetric, reflexive and transitive while representation is not.

Symmetry. If A is isomorphic to B then B is isomorphic to A (see above). But if A represents B, then B need not (and in fact in most cases does not) represent A. While a tube map represents the London Underground, the underground does not represent the map. Likewise, Maxwell's 'billiard ball model' represents a gas, but not vice versa.

One might now be tempted to counter that the relation between the billiard balls and the gas, for instance, is actually asymmetrical because a scientist has picked out features of the balls to represent the gas molecules. This is a good point, and I shall come back to this suggestion later on, but it is not available at present. The view I am

[17] The first two arguments I share in common with Mauricio Suárez with whom I have been discussing them over the last year. In essence, they have been put forward by Nelson Goodman (1968, 4-5) in connection with the similarity theory of pictorial representation.

discussing right now seeks to explain representation uniquely in terms of structural isomorphism and users have not yet entered the scene.

Reflexivity. Everything is isomorphic to itself, but most things do not represent themselves. To stick with the example, the map I hold in my hands represents the London Underground, but not the map itself.

Transitivity. Isomorphism is transitive but representation is not.[18] Botticelli's famous painting represents the birth of Venus. A photograph of if in an art history book represents the painting one can see in the *Galleria degli Uffizi* in Florence. But from this it does not follow that the photograph represents the birth of Venus. Suppose the gallery decides to sell the painting and a reproduction of it appears in Sotheby's auction catalogue. I take that what would be advertised for sale is Botticelli's painting, not Venus.

Chaos theory affords us with a telling example for the non-transitivity of representation in science. Most of the by-now famous mappings such as the cat-map, the horse-shoe, the tent map, or the baker's transformation have not been designed to model any real process. They are simplified and schematic pictures of an exceedingly complex continuous flow in an abstract phase space (which is in fact too complicated to be studied directly) and it is this flow that represents the dynamics of the real system. Hence, the mappings represent the phase flow and the phase flow represent the dynamics of the physical system, but the mapping does *not* represent the dynamics of the physical system.

[18] There is subtlety at this point (thanks to Mauricio Suárez for drawing my attention to this). To claim that representation is not transitive may amount either to claiming that it is *non-transitive* or that it is *intransitive.* A relation R is non-transitive if the following condition holds: Rxy & Ryz but not Rxz for some x, y, z. A relation R is atransitive if the following condition holds: Rxy & Ryz but not Rxz for all x, y, z. Non-transitivity is the weaker notion than intransitivity in that the latter implies the former but not vice versa. For the present argument the weaker notion will do: my claim is that representation is non-transitive.

(b) Isomorphism is not sufficient for representation

Structural isomorphism is too inclusive a concept to account for representation. In many cases neither one of a pair of isomorphic objects represents the other. Two copies of the same book, for instance, are perfectly isomorphic to one another but neither needs to be a representation of the other. Isomorphism between two items is not enough to establish the requisite relationship of representation; there are many cases of isomorphism where no representation is involved. Hence, isomorphism is not a sufficient condition for representation.

One might now argue that this critique is spurious since in the given set-up this problem cannot crop up at all.[19] The models under consideration are structures and the target systems are objects in the world. Counterexamples of the aforementioned type can then be ruled out simply by building this matter of fact into the definition of representation: introduce the ontological restriction that a model *M must* be a structure and the target *T must* be a concrete object (or process) in the world. This strategy blocks the above objection. Even if one book resembles the other, it does not represent it because it is the wrong kind of thing.

Unfortunately things are not that easy. Though models often do refer to things in the world, this is not necessarily so. Just as a picture can represent another picture, a model can represent another model rather than anything in the world. Consider again the aforementioned case of chaos theory. Mappings like the baker's transformation are models that represent what happens in another model, not what is going on in the world. But this is incompatible with the amendment suggested above since by requiring that the target of a model must be a real system we would rule out such obvious and important cases of representation. This is an unacceptable consequence.

(c) Multiple realisability

Looking at successful applications of mathematics in the sciences we find that quite often the same structure is used several times in different contexts and even across different disciplines. Linear spaces, for instance, are widely used in physics, economics, biology, and the mathmatised parts of psychology. Similarly, ordinal

[19] This argument parallels a suggestion discussed in Carroll (1999, 35-7) to rescue the similarity view of pictorial representation.

measurement scales are used to quantify length, volume, temperature, pressure, electrical resistance, hardness of a solid and many other things. The $1/r^2$ law of Newtonian gravity is also the 'mathematical skeleton' of Coulomb's law of electrostatic attraction and the weakening of sound or light as a function of the distance to the source. Harmonic oscillations are equally important in the context of classical mechanics and classical electrodynamics (for a detailed discussion of this case see Kroes 1989). These examples suggest that structures are 'one-over-many', as Shapiro (2000, 261) puts it. That is, the same structure can be exhibited by more than one target system. Borrowing a term from the philosophy of mind, one can say that structures are *multiply realisable.*

These examples are by no means isolated cases. Certain geometrical structures are possessed by many different systems. Just think about how many spherical things we find in the world; or consider the spiral we know from helter-skelters and spiral staircases and that later on even came to serve as a model of DNA. The same is true of many mathematical theories as well. Take once more the case of chaos theory. Many of its mathematical structures have been put to use in various different fields such as physics, biology, economics, medicine, or climate research. And similar remarks apply to other branches of mathematics such as partial differential equations or probability theory.

To see how the multiple realisability of structures clashes with the representational power of models, we need to bear the following feature of representations in mind. I observed at the outset that models are representations of *something else*. Implicit in this is that models are representations of some *particular* target system. The target can either be a token (as in the case of cosmological models) or a type (as in the case of models of the atom), but models are always representations of some physical phenomenon like an electric circuit, a falling object, magnetism in a solid, or an exploding star. Representations, by their very nature, are directed towards one particular phenomenon. This implies that the extension of a representation must be fixed correctly; a model of the hydrogen atom, say, represents hydrogen atoms and nothing else. Or to put it another way, the correct class of objects as its targets must be singled out for a scientific representation in much the same way as the objects a predicate or a proper name applies to must be singled out.

It is this feature of representation that is incompatible with multiple realisability. The point is that in the case of a multiply realisable structure, isomorphism is unable to single out the class of objects or phenomena that the model is intended to represent; that is, isomorphism cannot determine the correct extension of the representation. Or to put it another way, the model does not allow us to correctly identify the system it is supposed to represent. If a model is isomorphic to more than one system instantiating the same structure, which one is it a model of? What is the harmonic oscillator, to stick with the example, a model of? A pendulum bob? A lead ball on a spring moving up and down? The voltage in an electric circuit with a condenser and a solenoid? The amplitude of the B-field of an electromagnetic wave? The motion of atoms in the wall of a black body? It is isomorphic to all of them, but is it a model of all of them? This cannot be. To repeat, it is one of the main characteristics of a model is that it is a model of one particular phenomenon (no matter whether type or token). But if several parts of the external world can instantiate the same structure, a structure *per se* does not stand for one of them in particular. We face the dilemma that the structure *as a model* must stand for one particular system (or one particular type of system), but as a bare structure it is isomorphic to many systems (or types) and there is nothing that would allow us to pick out one of these as the 'privileged' one of which the structure 'really' is a model. In other words, the structure fails to indicate which one of the structurally isomorphic targets it is a model of.

Three counters to this point come to mind. The first tries to mitigate the force of the argument by denying that such cases are relevant to science. Van Fraassen (1980, 66), mentions a problem similar to the one at hand under the heading of 'unintended realisations' and then expresses confidence that it will 'disappear when we look at larger observable parts of the world'. Adopting his point to the present context one could try to counter the multiple realisability argument as follows: even if there are multiply realisable structures to begin with, they vanish as science progresses and considers more complex systems and larger chunks of the world. We have good reasons to believe that this solves the problem because the more we know about phenomena, the more likely it is that their structures will be different; and once we possess an accurate description no two phenomena will have the same structure.

There is a problem with this counter, however. It is just besides the point to appeal to *future* science to explain how models work *today*. It is a matter of fact that we currently have models that represent electric circuits and sound waves and a theory of scientific representation that cannot account for how they do so must be defective. We just do not have to await future science providing us with more detailed accounts of a phenomenon to make our models represent what they actually already do represent.

The second counter has it that the argument from multiple realisability is wrong because all that matters is that we can use a particular structure as a model for this or that. On one occasion we use the harmonic oscillator structure as a model for an electric circuit and on another occasion as a model for the motion of an atom. That is what scientists do all the time, so what is the problem with it?

There is no problem with *that*, but it describes a different set-up. This reply (at least implicitly) acknowledges that isomorphism by itself is not able to single out the relevant physical phenomena (or class thereof) and then appeals to users of the structure to accomplish this task. Though the appeal to users is a valid point, it is one that we cannot make at the present stage. Once again, the view I am discussing seeks to explain representation solely in terms of isomorphism and users are not part of that picture. To admit that something over and above isomorphism – e.g. users – is needed to pick out the right things in the world is to admit that (SM) fails.

The third counter is the most radical one in that it is an outright denial of the premise of the argument. There is no reason, so the counter goes, to assume that a model has to represent one particular phenomenon or that it must allow for an unambiguous identification of a particular phenomenon as its target. On this view, if a model is isomorphic to many different phenomena, then we have to bite the bullet and admit that it (simultaneously) represents all of them. Or to put it another way, what the model represents is not this or that phenomenon, but the class consisting of all phenomena whose structure is isomorphic to the one of the model.

This view is not impossible, but implausible. As a matter of fact, scientists do present us with models of phenomena and not structures. Take any science textbook at random and you will find models of the nucleus, models of galaxies, or models of predator-prey interaction, but nowhere you will find something like 'models of everything that has structure such and such'. To claim that all this is just idiosyncratic and superstitious jargon and that nothing except structures matters at the end of the day seems to stretch rational reconstruction beyond breaking point. Or to put it more mildly, someone wanting to take this line of argument bears the onus of proof.[20]

(d) Identity conditions for models

The last argument in this series is a kind of corollary to the above argument from multiple realisability. A successful account of models has to give us identity conditions for them, enabling us to say under what conditions two models are the same. That is, we must be able to individuate models (recall the slogan: 'no entity without identity'). But if models are taken to be structures that relate to reality via isomorphism this is not possible. The argument takes the form of a *reductio ad absurdum* and runs as follows. Let A and B be two target systems, different from each other, that instantiate the same structure (which is perfectly possible as we have seen in the previous paragraph) and let M_A and M_B be the respective models (i.e. M_A is a model of A and M_B is a model of B). Since by assumption A and B are different (they may be a pendulum bob and a electric circuit, for instance), their models must be different as well since models are about a particular system. However, since A and B instantiate the same structure and since, by assumption, models *are* structures M_A and M_B must the same. So we end up with the contradiction that M_A and M_B are and are not identical. Hence one of the premises must be false and I take it that it is the one that models are nothing but structures relating to reality via isomorphism. This shows that structures are not enough to individuate a model.

Where does representational power come from?

Throughout these arguments against (SM), there was a temptation to counter that an appeal to observers would make the problems vanish. This suggests that (SM) is overly 'purist' in stipulating that representation has to be accounted for *uniquely* in terms of isomorphism and that what we really need for representation is intention. Remedy then seems easy to get: concede that structures fail to be representations 'in themselves' and make users part of the picture. On such a view representations are

[20] Moreover, the argument in the next chapter is designed to show that this claim is wrong.

intentionally created; that is, structures become models when someone uses them as such.[21] From a formal perspective this amounts to making representation a triadic relation: a *user* takes *something* to be a model of *something else*. Working this into the structuralist view of models we obtain:

(SM') The structure S represents the target system T iff T is structurally isomorphic to S and S is intended by a user to represent T; given that this is the case, the model M consists of the structure S, the isomorphism that holds between S and T, and the user's intention to use S as a representation of T.[22]

At first glance, this appears to be a successful move, since in this version the above criticisms no longer go through. First, the appeal to intentions renders representation non-symmetric, non-reflexive and non-transitive. So the problem with the wrong logical properties vanishes. Second, the problem that isomorphism is not sufficient for representation is resolved 'by definition' since we simply stipulate that S represents. Third, the problem of multiple realisability vanishes because the user can intend S to be a representation of some particular system (e.g. an electric circuit) and forget about the other 'unintended' applications. Fourth, building intentions into the definition of the model undercuts the contradiction because M_A and M_B are no longer the same.

This move is so straightforward and so successful that it should make us suspicious. I agree that users are an essential part of an account of representation, but

[21] This is explicitly held by van Fraassen (1994, 170; 1997, 523 and 525).

[22] This view could be attributed to the German structuralists. Structuralists belonging to this camp emphasise that models are structures but at the same time recognise that structures *per se* do not stand for any particular bit of the external world. Structures only come to do so when an intended application is specified (see Balzer *et al.* 1987, Chs. 1 and 2). However, German structuralists are concerned with how different parts of a complex theory relate – that is, they are concerned with the architectonic of science, as the title of Balzer, Moulines and Sneed's book indicates – rather than with the functioning of an individual model. For this reason, they say comparatively little about the question of how an individual model faces reality and one may only guess what stance they would take on the issue.

merely adding the condition that someone intends to use S as a model of T to the standard structuralist account is not enough. There are several problems with that.

A first and simple reason why (SM') is not an adequate response to the enigma of representation is that isomorphism is not necessary for representation. As a matter of fact, many representations are inaccurate in one respect or another. The Bohr model of the atom or the liquid drop model of the nucleus, for instance, are models whose structure is not isomorphic to the structure of their respective target systems. On (SM') one would have to deny a model that involves some degree of inaccuracy the status of a scientific representation: either a model is isomorphic or it is not a representation at all. This, I take it, is too restrictive a view on scientific representation.

But (SM') is unsatisfactory for at least two more reasons. Fist, merely tacking on intentionality as an additional condition is question begging. Representation is an intentional notion, that much is granted. But what does that mean? I do not deny that scientists do make representations, and more specifically that it is they who turn something into a representation that would not be one otherwise. The question is how this happens. When we ask how representation works we want to know what exactly a scientist does when she uses S to represent T. If we are then told that she intends to represent T by means of S, this is a paraphrase of the problem rather than an answer because what we want to know is what this intention involves. What exactly do we do when we intend to use S as a representation? To make the thrust of this criticism clear, consider an analogue problem in the philosophy of language: by virtue of what does a word have the reference it does? We do not solve this problem by merely saying that a speaker intends words to refer to certain things. Of course they do, but this by itself does not answer the question. What we want to know is why the speaker achieves referring to something by using a word and coming to terms with this puzzle is what philosophers of language try to do in theories of reference. The situation in the philosophy of science exactly parallels the one in the philosophy of language: what we have to understand is why the scientist successfully comes to use S as a representation of T and to this end much more is needed than blunt appeal to intention. So we are back where we started.

Second, recall that the problem we are dealing with is the enigma of representation, the question why a model represents its target. Assume, for the sake of the argument, that (SM') provides us with a satisfactory response to this puzzle. If we then look at how the above-mentioned problems are resolved and at how the structure gets endowed with representational power, we realise that it is the appeal to intention that does all the work, while the original suggestion – that there has to be an isomorphism between structure and target – does not play any role at all in explaining why *S* represents *T*. One can replace isomorphism by any other formal requirement of how the two structures have to relate without changing anything in how representational power comes about. So isomorphism is just irrelevant to a response to the enigma of representation.

Some may feel uneasy about this and counter that something must wrong with this argument because isomorphism certainly is doing some work in the above account of how a structure is related to its target. Agreed, it does, but not the work we might think. Isomorphism regulates the way in which the model has to relate to the target. A model cannot stand for its target in any arbitrary way; it has to be isomorphic to it. Such a regulation is needed because an account of representation solely based on intention is too liberal. On such an account, nothing prevents us in principle from stipulating that the dot I have just put on a piece paper is a representation of a carbon atom. As long as representational power solely rests on intention, there is no way to rule out such cases. It is absurdities of that sort that are effectively undercut when isomorphism is added as a further requirement.

Stated this way, the problem becomes apparent: isomorphism, when used as above, is not put forward as a response to the enigma of depiction, but as one to the problem of *quomodity*. The function isomorphism performs within (SM') is to impose constraints on what kinds of representations are admissible. But as such it is a response to the problem of how a model represents, and not of why. So when insisting on isomorphism we have, presumably without taking notice of it, shifted the discussion from the enigma of representation to the problem of *quomodity*.

The bottom line is this: isomorphism is irrelevant to understanding how a model comes to represent something. Whether it is a sensible constraint to impose on the way in which a model represents is the subject of the next section.

3. Structuralism and the Problem of *Quomodity*

How does isomorphism fare as response to the problem of *quomodity*? My answer to this question will be sober as well. While isomorphism can be understood as one possible answer to the factual aspect of the problem, it is unacceptable as a normative stance.

The problem of *quomoditiy* in its factual variant is concerned with modes of representation. What different ways of representing a target are there? For sure, isomorphism is one possible general answer to this question. One way of representing a system is to come up with a model that is structurally isomorphic to it. This is an uncontroversial claim, I think, but also not a very strong one.

The emphasis many structuralists place on isomorphism suggests that they do not regard it merely as one way among others to represent something. What they seem to have in mind is the stronger, normative contention that a representation *must* be of that sort. In other words, they seem to think that it is a necessary condition for M to represent T that the two are isomorphic. The leading idea behind this claim is that only accurate representations count as representations and that isomorphism provides us with a criterion for what counts as accurate.

This is the strongest possible reading of the normative claim. I now argue that it is untenable and then discuss two weaker versions, which I also dismiss.

The claim that isomorphism is necessary for representation is mistaken for at least two reasons. First, as I have already mentioned, it is descriptively inadequate. Many representations are inaccurate in some way and as a consequence their structure is not isomorphic to the structure of their respective target systems. As a consequence, we would have to rule out cases of that sort as non-representational. Either the model is a representation and a fortiori isomorphic to its target, or it fails to represent altogether.

I take it that this is too restrictive. It is absurd to deny a model the status of a representation on the grounds that there is an inaccuracy in it. Scientific models almost always involve inaccuracies, and yet they are representations of some sort.

A second problem for this view is that it makes research a miracle. In order to assess the quality of a model, we have to presuppose that it is a representation. Only when we assume that M represents T can we ask the question of how well it fares.

We tentatively put forward a model as a representation of something and then try to find out how well it does its job. But this becomes impossible if we deny that an inaccurate model has representational power. False or inaccurate representations are also representations. On what grounds could we say that a model such as the flat disk model of the Earth is wrong if we deny its representational character? [23] To deny that false models represent undermines the process of testing a model.

Structuralists can counter that this is too strong a reading of their view. What they actually want to put forward is the much weaker claim that isomorphism is necessary not for representation *tout court*, but for *scientific* representation. To put it another way, the idea is that nothing less than accurate representation is good enough for science and that isomorphism provides us with a workable criterion of accuracy.

Unfortunately this view does not fare better than its bolder cousin. To deny the billiard ball model, say, the status of a scientific representation is almost as undesirable as denying it the status of a representation *tout court*. No one would seriously want to say that what scientists do when they tamper around with an inaccurate model falls outside the realm of science until they manage to restore isomorphism between their model and the target. On such a view hardly anyone would ever do science, if there would be any at all, which is absurd.

Therefore, isomorphism is not necessary for representation, neither scientific nor non-scientific.

The last, and probably most plausible reading of the claim that isomorphism is involved in scientific representation is that it is a regulative ideal. As science progresses, its models have to become isomorphic to their target systems. We may start with something overly simplistic or just blatantly wrong, but as we proceed we have to seek to establish isomorphism.

This claim, however, falls outside the realm of a theory of representation for it is nothing but convergent realism in structuralist guise. As such it belongs to the realism versus antirealism debate, which, as I pointed out in Chapter 1, should be kept separate from the issue of scientific representation. Of course, convergent realism is a time-honoured position one can hold, but as a view on representation it is

[23] For those who prefer less colloquial examples, think of Thomson's model of the atom (now commonly, in a rather pejorative manner, referred to as the 'pudding model') or the fluid model of heat conduction.

beside the point. Representations can be realistic, but they do not have to. Scientific modelling does not always amount to pointing a mirror towards things. So to make convergent realism a part of a theory of representation seems neither necessary nor desirable.

4. Amended Versions: Why They Do Not Fare Better

The leading idea of amended versions is to relax the requirement that model and target have to be isomorphic and use a less restrictive mapping instead. In this section I discuss three ramifications of the structuralist conception of models and explain why they do not fare better than the original suggestion.

Embedding. Redhead (2001, 79) points out that it is often too restrictive to require that all elements of a structure need to have corresponding bits in reality. Therefore, the appropriate relationship between model and target system is embedding and not isomorphism. From a technical viewpoint this simply amounts to changing the mapping f in the above definition of isomorphism slightly: stipulate that f is only injective and not necessarily surjective.[24]

Using embeddings instead of isomorphisms has the considerable advantage that the model can have 'surplus structure'. This is important to understand how some models work, in particular in modern physics where surplus structures abound. However, as far as the issue of representation is concerned, this move does not improve the situation. If a structure S_T is embedded in a larger structure S then S_T is isomorphic to a substructure S' of S. Moreover, it is the substructure S' that has representatitional power because, by definition, the 'rest' of S is surplus structure that does not correspond to anything in the world. But now we are back where we started: S' represents S_T due to their being isomorphic and therefore the problems the 'embedding version' of structuralism faces are identical to the ones of the 'isomorphism version'.

[24] A mapping f is injective iff: for all x, y, if x is not equal to y, then $f(x)$ is not equal to $f(y)$; it is surjective iff for any y in S there is a x in S_T such that $f(x)=y$ – or more colloquially, iff the mapping 'hits' the entire domain of S.

Homomorphism. We can further relax the isomorphism requirement by admitting mappings that are neither surjective nor injective. Formally, this provides us with the following definition: a homomorphism from S_T to S is a mapping f: $S_T \rightarrow S$ such that if the elements a_1, ..., a_n of S_T satisfy the relation R^T then the corresponding elements $b_1=f(a_1)$, ..., $b_n=f(a_n)$ in S satisfy R, where R is the relation in S corresponding to R^T (Mundy 1986, 395).

Not much ingenuity is needed to see that this does not take us very far. Homomorphism is as insufficient for representation as isomorphism: many things can be homomorphic without being representations of one another. The multiple realisation problem – and with it the problem with identity conditions – is even more severe as in the case of isomorphism. Since S_T and S need not have the same cardinality in order to be homomorphic, the class of structures that are homomorphic to a given structure S is much larger than the class of structures that are isomorphic to it. The only count on which homomorphism fares slightly better is the problem with logical properties: it is not symmetrical or transitive, though it is still reflexive. Finally, it is obvious that tacking on an intentionality condition – analogous to (SM') will not remove these difficulties.

As a response to the problem of *quomodity*, homorphism – once more – faces problems very similar to the ones of the isomorphism view. It is certainly a fair reply to the factual variant of the question, but not much else. Not all representations do conform to the pattern imposed by homomorphism (the same counter-instances can be invoked), and there is not reason why they should.

Partial Isomorphism. Over the last decade and a half Steven French and his co-workers have developed the concept of partial structures and partial isomorphisms (see references above). The main idea is to replace the relations in the structure by so-called partial relations, which allow for n-tuples of which we don't know whether they fall under the relation or not. More precisely, an n-ary partial relation R is one which is not necessarily defined for all n-tuples. That is, the n-tuples of elements of the domain of the structure divide into three mutually disjoint classes: the class R_1 consisting the n-tuples that fall under R, the class R_2 consisting of the ones that do not, and the class R_3 containing the ones for which it is not known (yet) whether they do or not. The last-mentioned class is necessarily empty if R is a 'standard' relation; it may be non-empty if R is partial. A structure that contains partial relations is referred to as a 'partial structure'. A partial isomorphism is a mapping f: $S_T \rightarrow S$ between two partial structures such that (i) f is one-to-one, (ii) if $(a_1, ..., a_n)$ is in R^T_1 then the corresponding tuple $(f(a_1), ..., f(a_n))$ is in R_1, and ditto for R^T_2 and R_2, where R is the relation in S corresponding to R^T in S_T. Obviously, we recover the standard definition of an isomorphism if R^T_3 and R_3 are both the empty set.

This manoeuvre, however, does not circumvent the problems of the isomorphism view. Partial isomorphisms still have the wrong logical properties; they are not sufficient for representation; partial structures are multiply realisable and give rise to trouble with identity conditions; etc.

Chapter 3

Further Strictures on Structures

1. Introduction

In order to make sense of the claim that there is an isomorphism between model and target, the latter has to possess a structure. Only structures can enter into an isomorphism relation and if the target were not structured, there would be no meaningful way to claim that it is isomorphic to the model. The assumption that target systems possess structures is not trivial. In this chapter I argue that a structure *S* can represent system *T* only with respect to a certain description. I present two independent arguments for this conclusion. First, I argue that the concept *possessing structure S* does not apply to a part of the physical world unless a more concrete description also applies. Therefore the claim that the target system possesses a certain structure is true only relative to the truth of a certain more concrete, non-structural description of the target system. This dependence carries over to isomorphism claims. If we claim that the target is isomorphic to some structure, this is true only relative to the truth of some more concrete description of the system. Second, I show that it is not true that a target system has one, and only one, structure. Depending on how we describe the system, it can exhibit different, non-isomorphic structures. As a consequence, descriptions cannot be omitted from an analysis of scientific representation and one has to recognise that scientific representation cannot be explained solely in terms of structures and isomorphisms.

2. Why One Cannot Have Structures All the Way Down

The argument: structures are not the whole story

In this section I argue that the concept *possessing structure S* is abstract relative to a set of more concrete, non-structural descriptions. As a consequence, isomorphism claims do not 'stand on their own' in that their truth rests on the truth of more concrete descriptions. To drive this point home, I first examine what kind of structures are employed in the context at hand and then show that they are abstract in the sense developed in Cartwright (1999), from which the conclusion follows.

Specific and unspecific structures

So far structures have been characterised as sets of individuals along with the relations in which they enter. But what exactly do we mean by 'individual' and 'relation', and hence by 'structure'? Upon closer examination we realise that there is an ambiguity in the use of these terms. On the one hand, the term 'structure' is used to refer to what I call 'specific structures', on the other hand to 'unspecific' structures.[25] A specific structure is, roughly speaking, an assemblage of certain concrete objects that exhibits a certain pattern. Bricks stacked on top of each other which 'relate' to make a wall, iron rods riveted together to build a bridge, neatly shelved books in a library, or the coloured stones of a mosaic are examples of concrete structures. Thus understood, 'structure' is closely tied to concrete entities. For many purposes in science, however, we want to be able to speak of two things having identical structure. So we need to employ 'structure' in a more general, i.e. 'unspecific', sense which allows us to say that, for example, a wall made of bricks and one made of stones have the same structure.

How can this be achieved? Simply by getting rid of the bricks and the stones! We reach the needed level of generality if we strip away from the structure everything that is 'material' because nothing about what the objects are matters for the definition of a structure – they may be whatever one likes them to be. As Russell

[25] This corresponds to Redhead's distinction between what he calls concrete and abstract structures (2001, 74-5). I use the terms 'specific' and 'unspecific' because I use 'concrete' and 'abstract' in a different sense later on.

puts it, the structure 'does not depend upon the particular terms that make up the field [i.e. the domain] of the relation. The field may be changed without changing the structure' (1919, 60). The only thing that matters from a structural point of view is that there are so and so many objects, be they bricks, rivets, stones or what have you, and that they are related to one another in a certain way. That is, we replace objects by dummies or placeholders; all we care about is that there are some things between which a certain relation holds, but we don't care about what they are.

A similar 'deflationary' move is needed in the case of the relations constitutive of the structure: it is not important what the relation 'in itself' is but only between which objects it holds. Russell again: '[f]or mathematical purposes [...] the only thing of importance about a relation is the cases in which it holds, not its intrinsic nature. Just as a class may be defined by various different but co-extensive concepts – e.g. "man" and "featherless biped" – so two relations which are conceptually different may hold in the same set of instances. [...] From the mathematical point of view, the only thing of importance about the relation "father" is that it defines this set of ordered couples.' (Russell 1919, 60) That is to say that we can specify relations purely extensionally – a relation is nothing over and above its extension, i.e. it is nothing but a class of ordered tuples.

This leaves us with dummy-objects between which purely extensionally defined relations hold; and this is all we have when we are dealing with an unspecific structure. To rigorous minds this may sound a bit obscure though. What exactly are these unspecific structures? As Redhead points out (2001, 75), there are two ways of characterising them. First, an unspecific structure can be thought of in an *ante rem* Platonic sense as an ideal form that is instantiated in all concrete objects that share the same pattern. These objects then form an isomorphism class. Second, those with less Platonic inclinations may adopt a more hardheaded approach and conceive of an unspecific structure just as the isomorphism class itself. Whichever of these two options one chooses, what we end up with is a notion of unspecific structures which takes them to be complex entities whose elements do not have any non-structural properties. That is, the elements of a structure have no properties save those that derive from their position in an extensionally defined pattern.

As an example consider the natural number structure. The zero of this structure has no other properties than those which follow from its being the zero of that structure – it does not posses weight, colour or any other property that is extrinsic to the structure. The essence of the natural numbers is their relation to other natural numbers, that is, the pattern common to any infinite collection of objects that is endowed with a successor relation, a unique initial object, and that satisfies the induction principle. Of course, this unspecific structure can be instantiated by many things, an infinite sequence of dots or an infinite sequence of distinct points in space, for instance, and Platonists would add: by abstract entities called numbers. But the fact that some (collections of) objects can instantiate the natural number structure does not add anything to the structure itself. The fact that a glass bead, say, takes the position of the number three in a certain set-up does not mean that the number three has the property of being transparent or hard. Being the number three amounts to no more and no less than being the third position in the natural number structure (compare Shapiro 2000, Ch. 10; Rickart 1995, 17; Dummett 1991, 295ff.).

Which type of structure is structuralism concerned with, specific or unspecific? It is obvious from the way in which structuralists make use of structures that what is at stake is unspecific structures. Redhead is explicit about this: 'Our claim will be that it is this abstract [i.e. unspecific] structure associated with physical reality that science aims, and to some extent succeeds, to uncover [...]' (2001, 75). In like manner, van Fraassen posits that '[a] scientific theory gives us a family of models to represent the phenomena [...] These models are *mathematical entiti*es, so all they have is structure [...]' (1997, 528-9, emphasis added; compare 522); and similarly: 'models [are] abstract structures studied in mathematics, which the theory advances as representations of these phenomena. [...] Since those models considered in their own right are mathematical structures, they are known only in the way things are known mathematically.' (van Fraassen 2001, 33). French and Ladyman (1999, 109), finally, affirm that 'the specific material of the models is irrelevant; rather it is the structural representation [...] which is important'.[26]

[26] Some decades ago, Wilfrid Sellers made a similar point when he argued against Mary Hesse that it is only structures of this sort (he refers to them as 'second order properties') that matter to the functioning of a model (1965, 180-84).

Different notions of abstraction

The term 'abstract' (originating from the Latin 'abstrahere' meaning to draw away, pull away, take away, take off, to strip, or to skin) is used in many diverging senses. Since abstraction is crucial to my argument, precision on the notion of abstraction at work is imperative. For this reason I here present a brief review of some important conceptions of abstraction that I do *not* use and then proceed to explicate the notion of abstract concepts on which my argument will be based.

One sense of 'abstract' is non-physical. At the most basic level this means that abstract objects are not accessible to perception; that is they cannot be seen, felt, smelled, tasted, or heard. It is in this way that philosophers of mathematics use the term when they say that numbers (and other objects of pure mathematics) are abstract. This characterisation, however, is unsatisfactory because of its relativity to human sense organs and its reliance on the unexplained notion of perception. How can these faults be fixed? On this opinions diverge. Not being in space and time, not being causally efficacious, or not being capable of undergoing intrinsic change, are just some of the criteria that have been suggested to demarcate the abstract from the concrete.[27]

Second, within the context of empirical science, abstraction is frequently construed as a process whereby we 'strip away', in our imagination, all properties from a concrete object that we believe are not relevant to the problem at hand. As a result we obtain a model in which only some of the potentially many factors or parameters are present, while the others are set aside as irrelevant. This allows us to focus on a limited set of properties 'in isolation'. If, for instance, we are interested in the shape of a triangular object, we disregard its temperature, colour, weight, hardness, and so on and just deal with its geometrical features. This process of stripping away is now commonly referred to as 'Aristotelian abstraction'.[28]

A third way of using the term 'abstract' goes back to Duhem (1906) who used it to characterise mathematical physics. According to Duhem, physics employs the simple and precise concepts of mathematics to represent complex and imperfect real

[27] For a fairly comprehensive list of criteria see Hale (1988, 86-87).

[28] See Cartwright (1989, 197) or Chakravartty (2001, 327-8). It is worth noting that Peirce discussed this type of abstraction at length under the heading of 'precissive abstraction', see Short (1988, 51), and Zeman (1982, 212-13) and references therein.

properties and for this reason physics is never able to adequately picture the world, not even in approximation. The concepts of physics are precise, but nature itself does not have this precise quantitative character. For instance, physics treats the Sun as perfect sphere, but the real Sun has an irregular surface that is not even solid and hence deviates considerably from its geometrical idealisation. And this is typical for modern physics: whenever we treat a problem mathematically, we end up using concepts that no longer describe reality. Mathematical physics fails to give us a realistic representation of the world and for this reason its concepts are abstract; or as Cartwright and Mendell put it: Duhem uses the notion of abstractness to separate theoretical from practical facts (1984, 137).

It may well be, and probably is, the case that structures are abstract in one or the other of these senses, but this is not what is doing the work in what follows. Central to my argument in this section is the notion of abstract concepts as developed in Nancy Cartwright's *Dappled World* (1999).[29] I will discuss this notion of abstraction in the next subsection and then point out some consequences.

Abstract Concepts

Some concepts are more abstract than others. *Game* is more abstract than *chess* or *soccer*; *work* is more abstract than *weeding the garden* or *cleaning the kitchen*; *enjoyment* is more abstract than *seeing a good movie*; and *travelling* is more abstract than *sitting in the train* or *riding a bicycle*. Intuitively it is quite clear why this is so. But what is it for one concept to be more abstract than another? Cartwright (1999, 39) provides us with two jointly necessary and sufficient conditions for a concept to be abstract:[30]

> 'First, a concept that is abstract relative to another more concrete set of descriptions never applies unless one of the more concrete descriptions also

---

[29] The use of this notion in what follows does not presuppose any of her views on other issues.

[30] Cartwright only takes these conditions to be necessary and remains silent about whether they are also sufficient. This caution, I suspect, stems from the fact that she is careful to set off her notion of abstractness from supervenience on the one hand, and the determinable-determinate distinction on the other. Whatever one takes as sufficient should distinguish among these three. However, for my argument nothing hinges on this difference and so I take these conditions to be sufficient as well.

applies. These are the descriptions that can be used to "fit out" the abstract description on any given occasion. Second, satisfying the associated concrete description that applies on a particular occasion is what satisfying the abstract description consists in on that occasion.'

To get a better grip on what these conditions amount to, consider the example of *travelling*. The first condition says that unless I either sit in the train, ride a bicycle, drive a car, or pursue some other activity that brings me from one place to another I am not travelling. The second condition says that my riding a bicycle right now *is* what my travelling consists in. The salient point is that I am only doing one thing, namely riding a bicycle. It is not the case that I am doing two things, riding a bicycle and travelling. I am doing just *one* thing since riding a bicycle is how I am travelling at this particular moment. To say that I am travelling is just a more abstract description of the very same activity.[31]

Let's render this a bit more precise. To start with, note that the relation 'being more abstract than' holds between two concepts; it holds, for instance between *game* and *chess*, or *game* and *pickup stick*. The concepts that are more concrete than one given concept form a family. The *game* example highlights an important feature of this family. The family of concrete concepts that fall under a more abstract one may be open in the sense that, first, we may not be able to enumerate all concepts that belong to it and that, second, this enumeration may change with time. No one is able to list all the games that exist at a given time and new games are invented every year. For this reason I use the noncommittal term 'family' rather than 'set' to refer to the collection of concepts with respect to which another concept is abstract. This, however, does not pose any threat to the current account of abstraction. Nothing depends on how many members the family has (it may have just one), on whether the collection of concreta is finite or infinite, on whether it is completely specifiable, or

[31] By the way, being confused on the issue of abstraction can lead to rather unfortunate outcomes also in rather mundane contexts. A US government official once tried to defuse people's concerns about pollution by trying to convince them that 'it isn't pollution that's harming the environment – it's the impurities in our air and water that are doing it'. Whatever one's stance on environmental questions, *this* certainly is not a reason not to take action against pollution.

on whether or not it is stable over time. So in a somewhat more formal rendering, the current notion of abstraction amounts to the following:

> A concept a is more abstract than the members b_i of a family $B=\{b_1, b_2, ...\}$ of concepts, where $b_i \neq a$ for all i,[32] iff
>
> (A1) For a to apply it is necessary that at least one member of B applies.
>
> (A2) On any given occasion, the fact that b_i, say, applies is what the applying of a at the same occasion consists in.

Finally note that these conditions define a relation between two concepts, 'being more abstract than', and not a monadic property, 'being abstract'. In a derivative sense, one could call a concept 'concrete' iff it applies without any other more concrete concept having to apply on the same occasion as well; and one could call all concepts that are not concrete abstract. An example for concrete concept in this sense would be *being cobalt blue.* But these notions of concrete and abstract will not be used later on.

Cartwright develops the current notion of abstractness using the relationship between fables and their moral as a starting point, thereby drawing on the theory of fables of G. E. Lessing, the great dramatist of the German Enlightenment. For this reason, having at least a brief look at the relationship between fables and their moral sheds some light on the present notion of abstraction. The leading idea is that the relationship between the moral and the fable is the one between the abstract and the concrete; or paraphrasing Lessing: a fable is a way of providing graspable, intuitive content for abstract, symbolic judgement (Cartwright 1999, 37-8). How does this work? Consider the following fable by Aesop 'A marten eats the grouse; A fox throttles the marten; the tooth of the wolf, the fox.' (*ibid.*, 39) and the general moral is 'the weaker is always prey to the stronger' (*ibid.*, 37). The relationship between the fable and the moral is that of the abstract to the more concrete. Here is why: 'The marten is wily and quick; the grouse is slow and innocent. That is what it is for the grouse to be weaker than the marten. The fox is weaker than the wolf. But this is not

[32] This clause is added to rule out that the relation *more abstract than* is reflexive, i.e. that every concept is more abstract than itself, which is absurd.

a new relation between the fox and the wolf beyond the ones we already know so well and can readily identify in the picture: the wolf is bigger, stronger, and has sharper teeth. That's what its being stronger than the fox consists in. [...]' (*ibid.*, 41)[33]

The abstractness of structure

In this subsection I come to the core of my case. I argue that the concept *possessing structure S* is abstract in the sense outlined above and that therefore it does not apply without some more concrete concepts applying as well.

What is needed for something to have a certain structure is that it consists of a bunch of discernible individuals (which together make up the domain of the structure) and that there are certain relations in which these individuals enter.[34] More precisely:

> A target system T possesses the structure $S=<U, R>$ iff T consists of individuals t_i such that:
>
> (1) there is a one-to-one mapping from $U=\{u_1, u_2, ...\}$ to $U_T=\{t_1, t_2, ...\}$ mapping u_i onto t_i for every i;
>
> (2) for all relations r_i of R and for all ordered tuples o_u one can form of elements u_i of U: $r_i(o_u)$ iff $r_i(o_t)$, where o_t is the tuple we obtain when replace every u_i by t_i in o_u.

Hence, to assert that a certain part of the real world (a target system) has this or that structure amounts to establishing that it consists of certain individuals which enter into a specific relational pattern. Trivially, this implies that some 'constituents' of the

[33] Note that, in some cases at least, there seems to be similar dependence running the other way as well (thanks to Julian Reiss for pointing this out to me). That is, some concrete concepts seem not to be applicable without some abstract concepts being applicable as well. For instance, it cannot be true of something that it is green without also being true of it that it has colour. Be this as it may, the only thing that matters to my argument is that one cannot have the abstract without the concrete. Nothing hinges on whether a similar dependence also runs the other way.

[34] Again, I omit operations here and in what follows. A justification for this omission has been given in Chapter 2.

system are individuals and others relations of this or that type, or to put it in a different way, that the concepts of *being an individual* applies to some parts of the system and *being a relation of type x* to others (where 'type x' is a placeholder for a formal characterisation of the relation specifying, for instance, that it is transitive or antisymmetric). The crucial thing to realise at this point is, I maintain, that *individual* and *relation of type x* are abstract concepts on the model of *enjoyment*, *work*, *game*, or *travel*. To call something an individual or a relation is an abstract assertion relative to more concrete claims, and if it is true then it is true relative to more concrete truisms.

Consider *being an individual*. We cannot apply this concept in the same way we apply *cobalt blue*, say. Even on the most basic understanding of *individual,* its applicability depends on whether other concepts apply as well. What these concepts are depends on contextual factors and the kinds of things we are dealing with (physical objects, persons, social units, ...). But this does not matter; the salient point is that whatever the circumstances, there are *some* notions that have to apply in order for something to be an individual.

To begin with, consider ordinary medium-size physical things. A minimal condition for such a thing to be an individual is that it occupies a certain space-time region. For this to be that case it must have a surface with a shape that sets it off from its environment. This surface in turn is defined by properties such as impenetrability, visibility to the human eye, having a certain texture, etc. If none of this is the case, we would not call a things an individual, at least not in the sense of an ordinary medium-size physical thing. If we change scale, other properties may become relevant; for instance having mass, charge, spin, or whatever property we take to be sufficient to be indicative of an individual. But in principle nothing changes: we need certain more concrete properties to obtain in order for something to be an individual. If something is neither visible nor possesses mass, shape, charge, nor any other identifiable feature, then it cannot be treated as an individual. And this also remains true outside the realm of physics, although matters may get more involved there.

And similarly in the case of *being a relation of type x*. Within the structuralist framework, relations are defined purely extensionally, i.e. as classes of ordered tuples, and have no properties other than those that derive from this extensional characterisation (i.e. transitivity, reflexivity, symmetry, etc.). I maintain that relations of this kind are abstract in the same way *pleasure* or *work* are. Take *being a transitive relation*, for instance. There are many transitive relations: *taller than*, *older than*, *hotter than*, *heavier than*, *stronger than*, *more expensive than*, *more recent than* (and their respective converses: *smaller than*, *younger than*, etc.), and with a little ingenuity one can extend this list *ad libitum*. By itself, there is nothing worrying about that. However, what we have to realise is that *being a transitive relation* is true of something only if either *greater than*, or *older than*, or ... is true of it as well. We cannot have the former without the latter; that is, something cannot be a transitive relation without also being one of the above listed relations. *Being taller than*, say, is what *being a transitive relation* consists in on a particular occasion. There simply is no such thing in the physical world as a relation that is nothing but transitive. In this *being a transitive relation* is like *pleasure*, if I neither see a movie, nor read a good book, nor ... I do not have pleasure and pursuing one of these activities is what my having pleasure amounts to at a given instant.

It is comforting to notice that all this is in line with how people who apply structures in the empirical sciences think about the subject matter. As an example consider the attribution of an ordinal measurement scale to a set of objects as discussed in Krantz et al. (1971). They define the task to be performed as 'assigning numbers to objects or events on the basis of qualitative observation of attributes' (*ibid.*, 1-2) and take the objects to be rods of some sort, which are assumed to be rigid and straight. Given this, we take two rods, a and b, place them side by side and adjust them so that they coincide at one end. As a result, either a extends beyond b at the other end, or b beyond a, or they both coincide. We then say, respectively, that a is longer than b ($a \succ b$), b is longer than a ($b \succ a$), or that a and b are equivalent in length ($a \sim b$) (*ibid.*, 2). After having done this, we assign numbers $\phi(a)$ and $\phi(b)$ to the rods in such a way that they reflect the results of these comparisons. That is, we require that the numbers be assigned so that $\phi(a) > \phi(b)$ if and only if $a \succ b$ (*ibid.*, p.2), where '$>$' is a sharp total ordering, i.e. a relation that is transitive and satisfies trichotomy (for details see Machover 1996, 33). By now it seems clear what I am getting at: what this simple example illustrates is that there is no way to apply an abstract concept like *being a sharp total ordering* directly, that is, without the 'mediation' of more concrete concepts. For it to be true of two rods that they stand in a sharp total ordering it must also be true of them that they are rigid, straight and that one extends beyond the other. On this occasion, *extending beyond* is what *being a sharp total ordering* consists in – if the former fails, the latter fails as well.

This is not the end of the story yet. Many of the more concrete relations (*greater than*, or *stronger than*, etc.) are themselves abstract in that they require other conditions to fall in place in order to be instantiated. Just think of Cartwritght's example of *weaker than* (discussed above), where the applicability of the concept depends on various concrete facts about the animals involved, such as *having bigger teeth than* or *having stronger claws than.* Once we get the gist of this example, it readily carries over to other contexts.

There are no general rules to decide at what point this 'backing' of abstract concepts by more concrete ones comes to a halt – if it does come to a halt at all. In some cases there may be many layers, in others just one (as in the above case of the ordinal measurement scale). But this is of no importance to the issue at stake. What matters is that the applicability of extensionally defined relations requires at least one layer of more concrete concepts.

Conclusion

The concept *possessing structure S* does not apply unless some more concrete concepts, which fit out the abstract descriptions of individuals and relations that occur in *possessing structure S* also apply. That is, *possessing structure S* only applies relative to a more concrete description of the target system. I refer to this description as the 'concrete description associated with *S*'. Therefore, the claim that the target system T possesses a certain structure S_T is true only relative to the truth of the claim that a certain more concrete description applies to T as well. For instance, the claim that T has structure S_T=[U={a, b, c}, R={(a, b), (b, c), (a, c)}] – the structure consisting of a three object domain endowed with a transitive relation – is true only relative to a more concrete claim of the following sort: T consists of three iron rods a, b, and c, where b is greater than a, and c is greater than both b and a.[35] Note that the concrete claims are not unique. There are many concrete claims that

[35] This is in line with how Suppes seems to think about this issue when he remarks that we get from a set theoretical model to a concrete model, as physicists conceive it, simply by thinking of the sets as sets of concrete objects (1970, Ch.2 pp. 8-9; see also 1960a, 13-14).

can make an abstract claim true. The salient point is that there has to be *some* concrete description that is true in order for the abstract description to be true because satisfying the associated concrete description is what satisfying the abstract description consists in some particular situation.

This dependence on more concrete truisms carries over to isomorphism claims. Isomorphism is a relation that holds between two structures. So if we claim that the target T is isomorphic to Structure S then trivially we assume that T has a structure S_T, which enters into the isomorphism with S. This in turn implies that there must be some more concrete description that is true of the system. Therefore, isomorphism claims are not 'primitive' in that their truth depends on the truth of a more concrete description.

This is not to say that we need these more concrete descriptions to study structures. Of course we can consider a bunch of objects along with transitive relation, say, and then see what features such a structure has without thinking about any concrete description at all. This is what mathematicians and logicians do. But we cannot do that once we want to use the structure to represent a certain part of the empirical world. We then have to claim that this structure is isomorphic to some target system and this involves claims about more concrete features.

Possible responses

Two reactions to this are possible. On the one hand, a moderate structuralist can point out that neither (SM) nor (SM') are committed to the view that isomorphism claims are primitive in that they do not rest on other truisms. On this moderate understanding of structuralism, the position only claims that there is an isomorphism between model and target but not that there is nothing over and above this. From this point of view, then, the above observation that isomorphism claims can be true only relative to certain more concrete truisms need not be understood as a threat to structuralism, or could even be welcomed as an friendly amendment.

On the other hand, a more radical brand of structuralism might want to resist this reconciliation and insist that the structuralist view of models, at least implicitly, is committed to the claim that nothing over and above structures and isomorphisms is needed to account for the representational function of models. This view needs to find a way around the above conclusion.[36] Two possibilities to do so come to mind. I discuss them in turn and conclude that they ultimately fail.

First, one might argue that the more concrete concepts that I argue are needed in order to make structural concepts applicable can themselves be analysed in structural terms. For this reason, so the argument goes, nothing non-structural is needed.

I cannot see how this could be possible. How do we analyse concepts like *having big teeth* or *being cobalt blue* in structural terms? In this case, I think, the burden of proof lies on the side of the structuralist and until a structural analysis of properties like the aforementioned is forthcoming I assume that this is not possible.

A second, more promising, line of argument may be suggested by van Fraassen (1997, 522-3). The problem he discusses is that if one structure can represent a phenomenon, then any other isomorphic structure can as well. For instance, we can represent temporal sequences either by using *greater or equal than* or *less or equal than*, but this does not, as he points out, mean that if *A* precedes *B* then *B* also precedes *A*. His solution to this problem is that we 'distinguish the two relations of less and greater – or more generally two isomorphic models – by noting their relations to each other in a larger structure in which they appear [...]' (1997, 523), This is to say, to put it in a slightly more technical way, that this ambiguity is removed by embedding the 'greater than structure' in a larger structure.

There are two ways of understanding this suggestion. On the first one (which is probably how van Fraassen himself would understand the problem), the question really only is how to distinguish between *greater or equal than* and *less or equal than*. His answer then is that the difference consists in yet more structural features. However, thus understood his suggestion, correct or not, is of no help in dissolving the above problem. Also a larger structure is just a structure and as such it is subject to the same objection: it is abstract and to assert that a part of the physical world possesses this structure is true only relative to more concrete truisms.

There is a second reading of this suggestion, which is probably not one that van Fraassen himself would adopt, but for the sake of argument it is still worthwhile to see where it takes us. This reading departs from the observation that, from a

[36] This seems to be the view of Da Costa and French when they argue that seemingly non-structural models such as icons and analogues can ultimately be analysed in purely structural terms (1999, pp. 258-63).

structuralist point of view, *greater or equal than* or *less or equal than* are in fact the same relation: they are blunt total orderings (i.e. they exhibit connectedness, weak antisymmetry and transitivity). This is literally all one can say about these relations from a structuralist point of view.[37] When van Fraassen paraphrases these relations as *greater or equal than* or *less or equal than* he has already moved one level down in the hierarchy of abstraction. Hence, one might take this proposal to amount to the claim that an abstract concept such *as being a blunt total ordering* turns into something more concrete (like *greater or equal than* or *older than*) by embedding it into a larger, more encompassing structure.

I don't think this is true. First, empirical claims have to stand on their own. There are many sciences that make extended use of mathematical modelling but which do not have, unlike some branches of physics, overarching theories. Population dynamics or economics are points in case. There simply is no larger structure in which a predator-prey model, say, could be embedded, but nevertheless we know what its relations express and how its terms connect to reality. Making empirical significance dependent on embeddability in a larger structure renders many good and fruitful models meaningless. This is a consequence, I take it, we are not willing to accept. Second, and more importantly, I just cannot see how the process described above is supposed to take place. Why should the fact that we embed one structure in another one endow it with empirical content it did not have before? Merely embedding some structure in a larger structure does not turn a transitive relation, say, into *greater or equal than*. To say that embedding solves the problem is just a non sequitur; sheer structure remains sheer structure, embedding something with no empirical content into something that has no empirical content does not yield any empirical content – there is no *deus ex machina* creating content out of nothing.

For these reasons, I think, the conclusion is inevitable: *Possessing structure S* is abstract on the model of *pleasure* or *work* and therefore the truth of an isomorphism claim rests on the truth of a more concrete description of the target system.

[37] Compare van Fraassen (1997, 516) where he points out that from a structuralist point of view '[s]cience is [...] interpreted as saying that the entities stand in relations which are transitive, reflexive, etc. but as giving no further clue as what those relations are.'

A qualification: a hierarchy of structures – but not all the way down!

At this point a qualification is needed. The above is an answer to the question of how a structure is connected to physical reality. But not all structures used in science are applied to reality *directly*. It would be wrong to think, for instance, that the structure described by Newton's equation of motion in its general form faces reality straight away. It is just after the equation has been 'adapted' to the concrete problem at hand (by specifying a particular force function and fixing certain boundary conditions) that it faces reality. Within the semantic view of theories, this fact is often accounted for by positing a hierarchy of structures with the most theoretical ones at the top and the most concrete ones at the bottom, which are connected via embeddings or other kinds of mappings. Another approach has been developed within the German structuralist tradition, where the notion of a theory net is invoked to account for how more general and more specific parts of a theory relate to one another (Balzer et al. 1987, Ch. 3). On this view, a theory is organised as an inverted tree-like net, whose knots are theory elements (e.g. equations or laws). The top-down arrangement reflects the relation of specification; that is, it reflects how central a certain element is. On that picture, Newton's equation would be at the top and the equation for a one-dimensional, linear, frictionless point-particle oscillator at the bottom. I will discuss this in the case of Newtonian mechanics in some detail in Chapter 8.

I am not at present taking a stance on how different parts of a complex theory relate. I am mentioning this issue only to point out that there is no contradiction between a view positing a layer-cake like set-up of structures (the semanticists' hierarchies or the German structuralists' theory nets, for instance) and my view of how structures connect to reality. If we stick with the layer-cake metaphor, I have only been concerned with the bottom level. Nothing in my view on what it means to say that a physical system has a certain structure commits me to the claim that *every* structure faces reality directly. There can be any number of layers whose relation is purely mathematical in nature. But, and this is the salient point, this does not get us all the way down. At some point the cake needs to stand on the table, as it were. Every cascade of structures has a bottom element and this bottom element needs to be anchored in reality in a way that is very different from how one would connect it to yet another structure. It is at this point at which the above considerations about more concrete descriptions become relevant. When we reach the bottom of the cascade we have to connect the 'last' structure to physical reality with the aid of a description of the by now familiar kind. This said, we get the following picture.

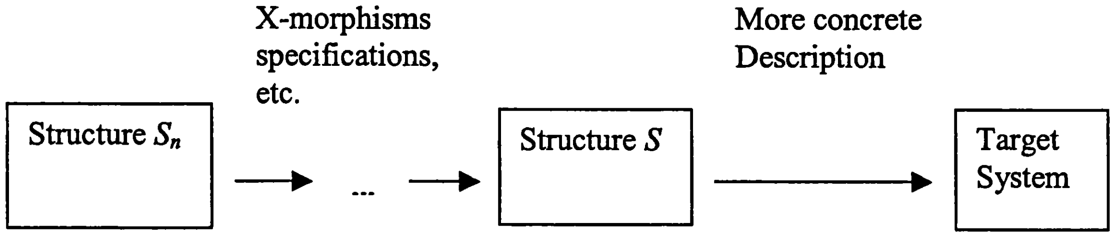

Fig. 2: How structures S_1, ..., S_p, where p is an arbitrary integer, face reality.

3. The Chimera of the One and Only Structure of Reality

There is not such thing as the one and only structure of a target system

The main contention of this section is that a target system does not have a unique structure. Depending on the conceptualisation we choose, the same target system can exhibit different, non-isomorphic structures. I first detail this claim, briefly explain how it differs from the no-go charge held against structuralism by Newman's theorem, and then provide examples illustrating the non-uniqueness of structure.

It is a necessary condition for something to have a structure that it consists of objects, which together make up the domain of the structure, along with relations in which these enter. So if a target system is to have a structure it has to be made up of objects and relations. And herein lies the rub. A piece of the physical world does not come sliced up with the pieces bearing labels saying 'this is an object' or 'this is a relation of this or that sort'. To cut up the system is a task that the scientist has to accomplish. This is important because what we recognise as individuals and what relations hold between these is, in part at least, a matter of how we conceptualise the target system. Structures result from a certain way of taking or demarcating the system; what the constitutive objects and relations in the target are depends on a schema of 'cutting up' the system. But different schemes may impose different orders on a domain. Depending our interests, objectives and standards we may choose different schemas, resulting in different structures. So there is no such thing as the one and only structure of a target system. A system has a determinate structure only relative to a certain conceptualisation; and given that there are different equally legitimate conceptualisations of a system there are different (typically non-isomorphic) structures that can be said to be structures of the system. Admittedly, there are ways of 'cutting up' a system that seem fairly simple and 'natural', while others may be rather involved or even contrived. But what seems contrived from one angle may seem simple from another one and from the viewpoint of a theory of scientific representation any is as good as any other.

What do I have to offer to sustain this claim? I don't think that there is a hard and fast general proof, at least I cannot think of one. My argument is inductive, as it were. In the subsections after the next one I discuss examples from various contexts and show that as a matter of fact it is possible to structure one system in different, equally valid ways when using different assumptions. These examples are chosen such that the imposition of multiple structures only relies on very general features of the systems at stake (e.g. their geometrical shape). For this reason, it is easy to carry over the strategies used to other cases as well. From this I conclude that there is at least a vast class of systems for which my claims bear out, and that is all I need.

Before I proceed, I should emphasise that I do not maintain that there is no way in which reality is; that is, I am not defending internal realism or some kind of metaphysical antirealism. Whether such a position is true or not is not my concern at the moment. What I am arguing for is the much weaker claim that things do not have one, and only one, structure. Moreover, it is interesting to note that metaphysical realism and structural anti-essentialism of the sort I am advocating do not exclude each other. In fact, even if we grant that there are natural kinds, they do not determine a unique structure. This is shown in the example below with the methane molecule which, though it is built up from hydrogen and carbon atoms (which are commonly taken to be natural kinds), does not have a unique structure.

Newman's theorem

In this subsection I make a few comments on Newman's theorem. This theorem has repeatedly been used as an argument against structuralism and since my points (seemingly) bear some similarities to it, it is worthwhile to point out what the differences are. In 1928, the Cambridge mathematician M.H.A. Newman proved a theorem stating (roughly) that any set can be structured in any way you like subject to cardinality constraints. In his own words (1928, 144): 'Any collection of things can be organised so as to have the structure *W*, provided there are the right number of them.' This sounds very much like my claim that systems may exhibit different, non-isomorphic structures. However, though the two arguments pull in the same direction in the sense that they aim at undermining a structural essentialism holding that every system has exactly one characteristic structure, there are important differences between the two.

My argument is, at once, stronger and weaker than Newman's. The proof of Newman's theorem turns on the fact that relations are understood extensionally in set theory (i.e. an n-place relation is taken to be no more than a set of ordered n-tuples). Hence, given a bunch of objects, we can structure them in any way we like just by putting the objects into ordered n-tuples; there are no constraints on this procedure other than that we need enough of them. No physical constraints are taken into account. This 'grouping together' of objects (which creates the relations and hence the structure) is a purely formal procedure that pays no attention to the nature of the objects involved. The relations so created do not need to have any physical reality (in the sense that they have any influence, observable or not, on the system's behaviour).

It is at this point that the two arguments diverge. What I want to argue is that a system can exhibit different *physically relevant* structures, i.e. structures that reflect the essence of a phenomenon and are not merely formal constructs. (I am aware of the fact that this is a somewhat vague characterisation but I that think there is no general description of what it means to capture the essence of a phenomenon. So I rely on the subsequent examples to clarify what I have in mind.) In this sense, my claim is stronger than Newman's. But at the same time it is weaker since the restriction to physically relevant structures drastically narrows down the freedom of choice. Among the great many structures compatible with cardinality constraints only a few will be physically relevant. Realistically, the scientist only has to choose between a limited number of structures – not anything goes. Which one of these we choose to use is often dependent on the context, the purpose of the investigation and other pragmatic factors. However, to discuss how these pragmatic factors determine certain choices is beyond the scope of this chapter. The point I want to argue for here only is that these alternative structures do exist.

A starting point: objects of everyday experience

To begin with, consider physical objects of our everyday experience. They are not presented to us in 'analysed' form; they do not normally consist of neatly defined parts. In most cases we face a 'compact entity' and we have to 'cut it up' before structures can be discussed. And this may not be an easy task. To get the flavour, consider a hum-drum example: what is the structure of the Eiffel Tower? We could, for instance, take all the rivets that keep the construction together as the objects and their spatial arrangement as relations. Or we could take the intersections of the iron bars as objects and the forces between them as relations. Or to make it simpler, we could take the three levels of the tower as individuals and consider them as structured by the relation 'higher than'. Or we could divide the tower into a North and a South part.

This small list by no means exhausts the possibilities, there are many more ways to carve the tower, and each may have its legitimacy. The first two may be of interest to construction engineers, the third to the provider of elevators, and the fourth to those who care about rust protection. The conclusion is: the object only has a structure relative to a certain way of conceptualising it. There is no such thing as *the* structure of the Eiffel Tower.

A continuation: the methane molecule

The methane molecule (CH_4) consists of one carbon and four hydrogen atoms. Since carbon and hydrogen have the same electro-negativity the four hydrogen atoms form a regular tetrahedron at the centre of which we find the carbon atom. If we now furthermore make the modelling assumption that atoms are point particles and assume the distances between these to be constant, then the space occupied by the molecule is a regular tetrahedron (which is to say that the compound has a tetrahedrical shape). In many scientific contexts (e.g. collisions or the behaviour of a molecule *vis a vis* a semipermeable membrane) only the shape of the molecule is relevant, that is, we can forget about the carbon atom sitting at the centre of the molecule. This naturally raises the question: What is the structure of (the shape of) the methane molecule?

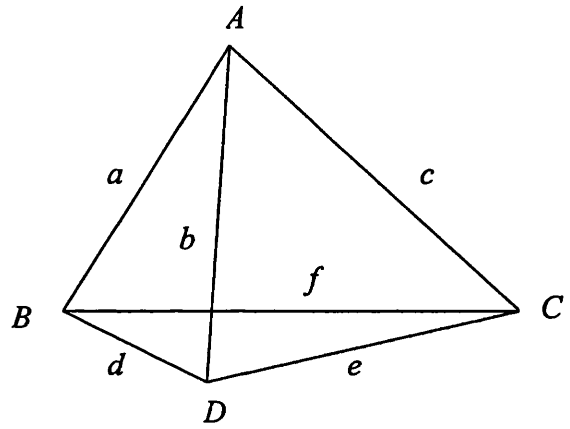

Fig 3. Tetrahedron.

To apply our notion of structure we need a set of basic objects and relations on it. Now the difficulties begin (see Rickart 1995, 23, 45). What are the objects that constitute the domain of the structure? A natural choice seems to take the corners (vertices) as the objects and the lines that connect the vertices (the edges of the tetrahedron) as the relations holding between the objects (the relation then is something like 'being connected by a line, or more formally Lxy = 'x is connected to y by a line'). We end up with the structure T_V which consists of a four-object domain, the four vertices A, B, C, and D, and the relation L which has the extension $\{(A, B), (A, C), (A, D), (B, C), (B, D), (C, D)\}$.

However, this is neither the only possible nor the only natural choice. Why not consider the lines as the objects and the vertices as the relations? There is nothing in the nature of vertices that makes them more 'object-like' than lines. Following this idea we obtain the structure T_S with a domain consisting of the six edges a, b, c, d, e and f on which the relation I (Ixy = 'x and y intersect') is defined. It has the following extension $\{(a, b), (a, c), (a, d)\ (a, f), (b, c), (b, d), (b, e), (c, e), (c, f), (d, f), (d, e)\}$.

So we need to conceptualise certain parts of the tetrahedron as objects and others as relations before we can tell what the structure of the (shape of the) molecule is. The tetrahedron *per se* has no structure at all; it just has a structure with respect to a certain description, namely one that specifies that the vertices are the individuals in the domain of the structure and the lines the relations, or vice versa.

But which one of these structures is *the* structure of the molecule? I think there is no fact of the matter. There is no reason why one structure should be privileged – both are based on 'natural' features of the tetrahedron. Hence, our molecule has (at least) two structures. And the important point is yet to come: These structures are not isomorphic. This follows immediately from the fact that the domains of T_S and T_V do not have the same cardinality.

This is not the end of the story yet. We can now replace one of the hydrogen atoms by another halogen, chloride say, and obtain CH_3Cl. Since chloride does not have the same electro-negativity as hydrogen, the new compound has the shape of an *irregular* tetrahedron, i.e. one whose edges do not have the same length. It is obvious that this tetrahedron still possesses T_S and T_V. However, due to its irregularity, it exhibits further structures as well: Take again the vertices to be the objects and define the relations between two vertices to be the lengths of the edges connecting them. We then get six different relations (assuming that all edges are of different length) whose extension comprises just one pair each. This new structure, call it T_L, is obviously non-isomorphic to both T_V and T_S. Hence we have ascribed three equally good structures to the irregular tetrahedrical shape of CH_3Cl just by adopting different modelling assumptions, and I don't doubt that with some ingenuity one can find many more.

Summing up, this example shows that molecules do not exhibit one particular structure in any obvious way. Objects of this sort may be analysed in more that one way with respect to structure. And this is by no means a peculiarity of the above example. The argument only relies on very general geometrical features of the shape of a molecule and therefore can easily be carried over to any kind of object, like cubes or other polyhedrons, or more generally objects consisting of lines (not even necessarily straight) that intersect at certain points.

Another continuation: the solar system

Another straightforward example illustrating my claim is the solar system (I will discuss this case in detail in Chapters 7 and 8). What we are usually dealing with when investigating the solar system is a composite entity consisting of ten spinning tops of perfectly spherical shape with a spherical mass distribution. One of these spheres has a privileged status in that it posses almost all the mass in the system. The other nine spheres are orbiting around this big sphere and interactions take place only between this big sphere and the small spheres, but not among them. The strength of the interaction, finally, is proportional to the inverse square of their distance.

For the purposes of celestial mechanics it no doubt is convenient to conceptualise the solar system in this way, but the choices made are by no means necessary. Why not consider the individual atoms in the system as basic entities? Or why not adopt a 'Polish' stance and also take the mereological sum of some planets as basic objects? And similarly for the relations between the objects: the choice to neglect all interactions but the ones between the Sun and the planets is by no means the only possible choice. There are many possibilities and it is the choices we make that finally demarcate the phenomena in a way that gives rise to a structure; it is just under this design that the system consists of clearly defined and identifiable parts and relations between them.

Logistic growth in ecological models

Suppose we are interested in the growth of a population of some particular species, wasps say. One of the earliest, and by now famous, models has it that the growth of the population is given by the so-called logistic map,

$$x' = Rx(1-x),$$

where x is the population density in one generation and x' in the next, R is the growth rate. For the sake of the argument, assume that this equation describes the situation correctly. From a structuralist point of view this amounts to saying that the structure that is defined by the logistic map, S_L, is isomorphic to the structure S_W of the system (i.e. the population of wasps) and that S_L represents the population of wasps.

A closer look at how the model really works quickly reveals that this is not the whole story. As Hofbauer und Sigmund (1998, 3) point out, in many ecosystems thousands of different species interact in complex patterns and even the interactions between two species can be quite complicated, involving the effects of seasonal variations, age structure, spatial distribution and the like. But nothing of this is visible in the model. No interaction with any part of the ecosystem is explicitly built into the model. It is just the net effect of all interactions that is accounted for in the last term of the equation ($-Rx^2$), which reflects the fact that a population cannot grow infinitely due to restrictions imposed by the environment. Hence, all actual interactions are 'idealised away'.

Once this is done, one has to define the objects of the structure. An obvious choice would be individual wasps. But one readily realises that this would lead to huge and intractable sets of equations. The 'smart' choice therefore is to take generations rather then individual insects as objects. But this is not enough yet. We have to assume furthermore that the generations are non-overlapping, reproduce at a constant average rate (reflected in the magnitude of R) and in equidistant discrete time steps.

The thrust of the story is clear by now: there is no structure just 'sitting there' waiting to be picked up. The analysis of certain structures into the target system is a quite difficult and laborious task and the given system could be analysed in different ways as well, depending upon which aspects of the system one wishes to emphasise. Before we do not decide to take generations as basic units of study, make time discrete, neglect the interactions with other parts of the ecosystem, and so on, one cannot meaningfully say that the target system exhibits one particular structure.

A final remark

I should emphasise that there is no temporal component to this sketch of how a particular structure is assigned to the target. The above is not meant to suggest that what we do when thinking in structural terms is *first* discern parts of a system and identify relations among these non-mathematical objects and *later* pick mathematical structures to represent them. In actual practice, it is often decided in advance what, or at least what kind of, structure is to be used and this choice to a great extent guides modelling assumptions on the more concrete level right from the beginning. In the above case of the population, for instance, it is obvious that the choice to model generations as discrete and equidistant in time and to assume a constant reproduction rate is tailored towards the application of iterative mappings.[38] But this does not threaten the above line of argument. What motivates certain modelling assumptions is unimportant to an account of how structures enter the scene. What matters is that at the end of the process of the construction of the model we are actually able to specify the concrete model we use and to account for how it relates to the structure we employ. How we get there is irrelevant.

4. Counter: This Is Nonsense – All We Need Is a Data Model

A possible counter to this is suggested in the writings of van Fraassen. This counter has it that all I have said so far is wrongheaded from beginning to end because it misconstrues the nature of the target system. I have assumed that what a model represents is an object (or event) of some sort. But, so the objection goes, this is mistaken. What a model ultimately represents is a data model, not an object. In this section I discuss this objection and explain why I think it is misguided.

Let me begin by introducing the concept of a data model (Suppes 1960b). Roughly speaking, a data model is a corrected, rectified, regimented, and in many instances idealised version of the data we gain from immediate observation, the so-called raw data. The data as provided by an experiment are often not in a form in which they are useful to scientists and they have to undergo processing before they can be analysed. This processing varies greatly in different scientific contexts. Characteristically, we first eliminate errors (e.g. remove points from the record that are due to faulty observation) and then present the data in a 'neat' way, for instance by drawing a smooth curve through a set of points. These two steps are commonly referred to as 'data reduction' and 'curve fitting'. As a simple example consider astronomical observation. What we take down on paper are the co-ordinates of a certain planet, say, at a certain instant of time. Given this record, we eliminate points

[38] Ketland (2000) even argues that we always use what he calls a background structure to 'slice up' a physical system.

that are fallacious and then fit a smooth curve to the remaining ones. As a result we obtain a data model of the motion of the planet.

A possible structuralist objection to the above arguments is that it is data models of the kind just described that are the targets of (theoretical) models, not towers, molecules or populations; that is, on this view the only thing a model represents is data:

> 'The whole point of having theoretical models is that they should fit the phenomena, that is, fit the models of data.' (van Fraassen 1981, 667)

> '[...] the theoretical models (proffered [...] as candidates for the representation of the phenomena) are confronted by the data models. [...] to fit those data models is ultimately the bottom line.' (van Fraassen 2002, 164)[39]

This move solves the problem in an elegant way. Data models are mathematical entities and as such can be considered to have a well-defined structure. This, so the objection goes, makes my entire argument obsolete.

In the remainder of this section, I will discuss this counter and say why I think it is wrong. But before that an exegetical remark needs to be made. The problem is that it does not become clear from van Fraassen's writings whether or not he thinks that the fitting relation between the model and the data *is* the representation relation or not, as the above counter assumes. On the one hand he repeatedly says that models are structures and that they represent observable phenomena and then goes on saying that these models fit the data (1980, 64; 1997, 524; 2001, 31; 2002, 164). This might suggest that representation *consists in* fitting the data; in short, that representation *is* data fitting. On the other hand, van Fraassen, to my knowledge, never explicitly states that this is so. He characteristically says something along the lines that models represent and their success consists in their fitting the data (*ibid.*). This, however, leaves it open whether the fitting relation *is* the representation relation or whether fitting is merely a criterion of empirical adequacy (and hence acceptability) without itself being the representation relation. If the latter, it is at least in principle possible

[39] Compare also van Fraassen (2002, 164, 252; 1997, 524; 2001, 31; 1989, 229; 1985, 271) and French (1999, 191-92).

that there is a representation relation between model and target that is distinct from the fitting relation. This leaves us with two possibilities: fitting is or is not equivalent to representation. If the former is the case, that is if fitting and representing are distinct, then we are back where we started because neither van Fraassen nor any other structuralist gives us any indication as to what this representation relation might be. If the latter is the case, we have a substantial suggestion as to what representation consists in. So it seems worthwhile to try out whether this is a workable suggestion, irrespective of whether van Fraassen would endorse it or not. This is the project for the remainder of this section.

This suggestion is wrong, I think, simply because it is descriptively inadequate. It is just not the case that models represent data. This point is not new. It has been argued at length by Jim Bogen and Jim Woodward (1988), Woodward (1989), and has recently been reiterated in different guise (and without reference to Bogen and Woodward) by Paul Teller (2001). In essence I agree with these authors and in what follows I have little to add to the substance of their arguments; therefore I will be rather brief. My focus, however, differs slightly from theirs and I therefore present the subject matter in a way that suits my needs.

The basic problem with the suggestion that what a model represents is a data model is that when we look at actual models this is not what happens. Most models do not *per se* contain anything that could be directly compared to data we gather; or more specifically, they do not involve structures that could plausibly be thought of as being isomorphic to a data model. This is of course not to say that a model cannot be compared to reality; the claim only is that this comparison does not happen in the way that the suggestion at hand has it. Let me detail this by dint of three examples, from which I then draw some general consequences.

To begin with, consider the discovery of weak neutral currents (Bogen and Woodward 1988, 315-18). What the model at stake consists of is particles: neutrinos, nucleons, the Z^0, and so on, along with the reactions that take place between them.[40] Nothing of that, however, shows in the relevant data. What was produced at the CERN were 290000 bubble chamber photographs of which roughly 100 were

[40] The model I am talking about here is not the so-called standard model of elementary particles as a whole. Rather, what I have in mind is one specific model about the interaction of certain particles of the kind one would find in a theoretical paper on this experiment.

considered to provide evidence for the existence of neutral currents. The notable point in this story is that there is no part of the model (which quantum field theory provides us with) that could be claimed to be isomorphic to these photographs (or any data model one might want to construct on the basis of these). It is weak neutral currents that occur in the model, but not any sort of data we gather in an experiment.

This is not to say that these data have nothing to do with the model. The model posits a certain number of particles and informs us about the way in which they interact both with each other and with their environment. Using this we can place them in a certain experimental context. The data we then gather in an experiment are the product of the elements of the model and of the way in which they operate in a given context. Characteristically this context is one which we are able to control and about which we have reliable knowledge (e.g. knowledge about detectors, accelerators, photographic plates and so on). Using this and the model we can derive predictions about what the outcomes of an experiment will be. But, and this is the salient point, these predictions involve the entire experimental set-up and not only the model and there is nothing in the model itself with which one could compare the data. Hence, data are highly contextual and there is a big gap between observable outcomes of experiments and anything one might call a substructure of a model of neutral currents.

To underwrite this claim notice that parallel to the research at CERN, the NAL also performed an experiment to detect weak neutral currents. The data obtained in this experiment were quite different, however. They consisted of records of patterns of discharge in electronic particle detectors. Though the experiments at CERN and at NAL were totally different and as a consequence the data gathered had nothing in common, they were meant to provide evidence for the same theoretical model. But the model, to reiterate the point, does not contain any of these contextual factors. It posits certain particles and their interaction with other particles, not how detectors work or what readings they show. That is, the model is not idiosyncratic to a special experimental context in the way the data are and therefore it is not surprising that they do not contain a substructure that is isomorphic to the data.

Before drawing some general conclusions from this, let me add two further examples, one from immunology and one from geology. This seems appropriate because the previous example may have left some under the impression that the discrepancy between what a model is about and what we observe is either peculiar to physics or is rooted in the fact that the items in the model are too small to be seen. The purpose of the next examples is to defuse this impression.

Next, look at the case of the infection with a virus, HIV say. An immunological model of the infection with HIV consists of some basic constituents of the immune system of a certain organism along with their interactions, and a mechanism of how the virus gets into the body, how it spreads, and how it reproduces. But all that is still highly theoretical. What do we observe? In the first instance nothing at all. It is only once the HIV infection develops into AIDS that certain other symptoms such as rashes, herpes, pneumonia, meningitis, tuberculosis and others become apparent. The other possibility for observing the infection is by conducting a test. But then, too, the evidence is rather indirect: we read off certain numbers from a machine for the analysis of blood or observe the change of colour in a test-tube, depending on what type of test we make. However, and this is the salient point, at no stage in this process is there anything that could be compared (let alone be claimed to be isomorphic) to the immunological model of infection (or parts thereof). In fact, nothing of what is observed in the laboratory is even considered to be part of immunology.

Lastly, consider geological models of plate tectonics (Giere 1988, Ch. 8). The basic entities occurring in models of that kind are magma, which is carried to the surface of the Earth by convection, and tectonic plates, which are carried along different divergent streams. The geological data, however, do not show any of these. What geologists do is not measure the speed of the moving plates directly, which would be a hopeless enterprise. They exploit the fact that the Earth possesses a magnetic field, which leaves its trace in the magma when it hardens, and which is known to reverse periodically. Given this, geologists observe magnetic properties of the hardened magma that we find at the edges of a plate and investigate its magnetic properties. As a result they obtain a description of the alternating patterns of magnetisation. The upshot of this is the same as above: the data collected are not isomorphic to anything in the model. If anything, it is magnetic patterns that we observe and that are reflected in the data (or the data model we construct from the raw data), but not the movement of tectonic plates.

The lesson to draw from these examples is the same as in the case of the weak neutral currents: there is nothing in these models *per se* that could plausibly be claimed to be isomorphic to the data we gather in observation. AIDS tests change frequently as research progresses and even at one given instant of time there are several different tests in use, which provide us with rather different sets of data. Nevertheless it is always the same immunological model that is at stake. Similarly, a geologist may collect quite different sets of data, yet the theoretical model they are compared with remains the same.

In line with Bogen and Woodward one can offer the following response to this: distinguish between phenomena and data and claim that models represent the former but not the latter. The phenomena represented in the above cases are weak neutral currents, the presence of the HIV virus in an organism and its effect on the immune system, and the movement of magma inside the Earth and the dynamics of tectonic plates. The data are bubble chamber photographs, readings on a blood analysis machine and records of the magnetisation of hardened magma.

It is difficult to give a general characterisation of phenomena because they do not belong to one of the traditional ontological categories (*ibid.*, 321). In fact, phenomena fall into many different established categories, including particular objects, features, events, processes, states, states of affairs, or they defy classification in these terms altogether. This, however, does not detract from the usefulness of the concept of a phenomenon because specifying one particular ontological category to which all phenomena belong is inessential to the purpose of this section. What matters to the problem at hand is the distinctive role they play in connection with representation.

What then is the significance of data, if they are not the kind of things that models represent? The answer to this question is that data perform an evidential function. That is, data play the role of evidence for the presence of certain phenomena. The fact that we find a certain pattern in a bubble chamber photograph is evidence for the existence of neutral currents; or the presence of a certain magnetic structure in hardened magma is evidence for the motion of tectonic plates. Thus construed, we do not denigrate the importance of data to science, but we do not have to require that data have to be isomorphically embeddable into the model at stake.

Needless to say, what counts as evidence for what is an involved question, but not one that needs to trouble us now.

The constructive empiricist might now reply that by postulating phenomena over and above data we left the firm ground of observable things and started engaging in fruitless speculation. But science has to restrict its claims to observables and remain silent (or at least agnostic) about the rest. Therefore, so the objection goes, phenomena are chimeras that cannot be part of any serious science.

Whatever stance one wants to take on this issue – in fact, one can construe phenomena in a realist (Bogen and Woodward 1988) as well as an antirealist (McAllister 1997) fashion – it won't help the structuralist. Denying the reality of phenomena is beside the point. Irrespective of whether one takes phenomena to be part of the furniture of the world, social constructions, or sheer aids for the economy of thought, it is these that models portray and not data. Or to put it more crudely: to deny the realty of phenomena just won't make a theoretical model represent data. Whether we regard neutral currents as real or not, it is neutral currents that are portrayed in a field-theoretical model, not bubble chamber photographs. Of course, one can suspend belief about the reality of these currents, but that is a different matter.

In conclusion, the suggestion at hand is untenable: representation cannot be explained in terms of an isomorphism between a data model and an empirical substructure of the model.

5. Conclusion: Structures and Descriptions Go in Tandem

In this chapter I argued, first, that the concept *possessing structure S* is abstract relative to a more concrete description and, second, that a target system can exhibit different structures relative to different descriptions and that for this reason there is no such thing as the one and only structure of a system. As a consequence, the structuralist notion that a structure S represents the target T iff the two are isomorphic needs to be qualified. In order to make sense of the claim that S and T are isomorphic we have to assume that T possesses a structure S_T, because only structures can enter into an isomorphism relation. But T possesses a structure only relative to a certain more concrete description. From this it follows that S represents a target system T only with respect to a certain description D; without such a description the claim that S is isomorphic to T is simply meaningless. Hence, descriptions are an integral part of any workable conception of scientific modelling and we cannot omit them from our analysis of representation.

This does not refute structuralism, but it puts it into perspective.[41] Nothing that has been said in this chapter renders the notion that a structure S represents its target T by dint of being isomorphic to it incoherent; but one has to recognise that isomorphism cannot be had without the mediation of a more concrete description. And this is more than a friendly, but slightly pedantic and ultimately insignificant amendment to the structuralist view. Every complete account of scientific representation has to explain how the interplay of structure, description, and reality works. If I am right on this, the face of discussions about scientific representation will have to change. The role descriptions play in scientific representation has not been acknowledged, let alone systematically discussed in current debates. So if the discussion in this chapter tells us one thing, then it is that we need to get descriptions back into the picture!

A sceptic might still reply that although there is nothing wrong with my claim that we need descriptions, there is not much of an issue here. What we are ultimately interested in, so the objection goes, is the isomorphism claim and that such a claim is made against the background of some description or another may be interesting to know, but without any further significance. I disagree. Phrases like 'S is isomorphic to T *with respect to* description D', 'S is isomorphic to T *relative to* description D', or 'isomorphism claims *operate against the background* of description D' point in the right direction and convey the leading idea; but they are delusive in that they might make us believe that we understand how the interplay between structure, world and description works. Nothing could be farther off than that. These expressions are too

[41] It is interesting to note that the German structuralists do explicitly acknowledge the need for a concrete description of the target system (Balzer *et al.* 1987, 37-8). Moreover, they consider these 'informal descriptions' to be 'internal' to the theory. Unfortunately they do not dwell on this issue and many questions (for instance concerning the character of these descriptions, the work they do within the architecture of the theory or how the interplay between abstract and more concrete descriptions work) are left unanswered.

vague to take us anywhere near something like an analysis of scientific representation. More needs to be said about how structures, targets and descriptions integrate into a consistent theory of representation.

The conclusion is inevitable: we have to go back to the drawing board. This is what I do in Part III. At this stage, I only briefly want to anticipate one element of the views on representation that I develop in the next part. This is in order to complete my criticism of the structuralist conception of models. So far I have just argued that we need to bring descriptions back into the picture. The problem is *how* to do this. In my view, the right way to do so is to integrate what descriptions describe into the unit we call 'the model'. (A cautionary note is needed here. I do not suggest making *descriptions themselves* part of the model. What I suggest – and what I argue for in detail in Chapter 5 – is that *what descriptions describe* has to be integrated into the model.)

That this is the right move is a result that follows naturally from the discussion in Chapter 5. But there are reasons independent of my views on modelling developed in that later chapter to think that what descriptions describe should be part of the model. In the remainder of this section I explain what these reasons are.

How we conceptualise a part of the world is an essential aspect of how we represent it. How we carve a system, how we 'see' it, and how we represent it are the two sides of the same coin. As the argument in the previous section shows, a part of the physical world only exhibits a well-defined structure relative to a certain conceptualisation and hence it can be related to other structures (for instance via isomorphism) only with respect to this conceptualisation. Given this, it is putting things upside down when we claim that the model is a structure, and nothing but a structure. In the light of the role descriptions play it seems only natural to make what they describe part of the model in one way or another.

One might try to resist this conclusion on the basis of an analogy with language. All kinds of things are needed to make a sentence represent a certain matter of fact. One may need causal chains connecting the terms in the sentence to their referents, one may need the speaker's intentions, one may need knowledge about the pragmatics involved, or what have you. But, so the objection goes, we would not want to make all this part of the sentence. In short, not everything that makes the sentence represent a certain matter of fact must be part of the sentence.

Two replies to this come to mind. First, one might deny that the analogy is legitimate. It may well be that models work in way that is very different from how sentences work and that what is true of sentences need not be true of models. Second, even if one buys into this analogy, there seems to be something murky about it. Quibbles about the difference between sentences and propositions aside, sentences have content. And more to the point, content *is* a part of the sentence. It would be absurd to take sentences to be no more than ink scribbles on paper and construe content as external to them.[42] And the same is true of models. Models have content – though they may have content in way that is different from how sentences have content – and their content is internal to them. Then, given what I said about the interplay between structures and descriptions in this chapter, it seems natural to say that what the description describes is part of the content of a scientific representation. If we then assume, as I do, that models are the vehicles of scientific representation, the conclusion follows: what a description is about is part of 'the model'. And therefore the radical structuralist notion that models are structures, and nothing but structures, is untenable. Of course, I by no means want to suggest that structures are not part of the picture. I only deny that structures are all that there is to scientific models.

[42] This is true of other kinds of representation as well. Mental representations have content. No one doubts that. The question is what their content is, how they acquire content, and how they 'encode' content. And the same goes for pictorial representations. It is an interesting question why the family photograph in front of me has content, but there is no doubt that it has.

Chapter 4

Putting Similarity into Perspective

1. Models and Similarity

According to an alternative version of the semantic view of theories, the relation between a model and its target is similarity (or resemblance, I use these interchangeably) rather than isomorphism.[43] As with the isomorphism view, the problem is that proponents of the similarity account do not explicitly address the issue of scientific representation. So it is not clear what their stance on representation is. I take it that a straightforward and fair account of representation in the spirit of a similarity view of models would be the following: the model M represents the target system T iff M is similar to (or resembles) T.

This view imposes fewer restrictions on what is acceptable as a scientific representation than the structuralist conception. First, it enjoys the considerable advantage over the isomorphism view that it allows for models that are only approximately the same as their targets. Given that most models involve inaccuracies of some sort or another, this is an important improvement. Second, the similarity view of modelling is not committed to a particular ontology. Although it is sometimes suggested that models are structures, the similarity view is not committed to this claim. Unlike the isomorphism view, it enjoys complete freedom in choosing its models to be whatever it wants them to be. The only restriction the similarity view

[43] The view that representation involves similarity can be traced at least to Plato and it has appeared in different contexts and in different guises throughout the history of philosophy (see Carroll 1999, Ch. 1 or Gordon 1997, Ch. 5 for surveys). In recent debates within the philosophy of science the similarity view of representation has been put forward most forcefully by Ronald Giere (1988, Ch. 3; 1999, 2002).

needs to impose on what models are is that they belong to the sort of things that can be similar to other things. But this is not much of a restriction, everything can.

In the light of its undeniable appeal and popularity, this view deserves careful consideration. But before embarking on a discussion, we have to get clear on what stance it takes on the three basic problems of a theory of representation. As I just observed, the similarity view is not committed to a particular ontology. This makes the task of this chapter easier: where there is no claim, there is no need for criticism. For this reason, the discussion in what follows is entirely concerned with the two semantic conundrums. Like the isomorphism view, it is ambiguous about whether similarity is supposed to be an answer to the enigma of representation or to the problem of *quomodity*. I will discuss each possibility in turn. My conclusion will be that as regards the enigma of representation the similarity view does not fare better than its structuralist cousin. As a response to the problem of *quomodity* it is an acceptable view, but one that is not very telling because the expression 'is similar to' is little more than a blank that needs to be filled in every instance in which it is invoked.

2. Similarity and the Enigma of Representation

Similarity does not fare better than isomorphism when understood as a response to the enigma of depiction. The problems it faces by and large parallel those of isomorphism. There is no point in repeating all the arguments put forward in Chapter 2 in detail and therefore I confine myself to indicating wherein the communalities lie and where, at times, the arguments diverge.

To begin with, similarity has the wrong logical properties: it is symmetrical and reflexive, representation is not. However, similarity fares better than isomorphism when it comes to transitivity. Similarity is not generally transitive and for this reason the objection from transitivity does not go through.

As with isomorphism, similarity is not sufficient for representation. Many things are similar without being representations of one another. And this is not a matter of the degree of resemblance. Even perfect similarity between two items is not enough to ensure that one of them represents the other. Two photographs of the same scene resemble each other perfectly, but they are not representations of each other.

The multiple realisability argument carries over to similarity as well, not in letter but in spirit. A particular model may be similar to several different things without being a representation of all of them. I observed that a model is a model of some particular physical phenomenon. This is incompatible with similarity being the principle by which representation works. In a case in which the model is similar to several different phenomena similarity fails to single out which phenomenon the model stands for. That is, similarity cannot determine the correct extension of the representation.

From this it becomes clear that the fourth argument against isomorphism, the contradiction that arises in connection with identity conditions, carries over to similarity *mutatis mutandis*.

One might now be tempted to say that an appeal to users solves the problem. Of course it won't, and this for exactly the same reasons as in the isomorphism case. Summing up, similarity does not do as a response to the enigma of depiction.

3. Similarity and the Problem of *Quomodity*

Can similarity be understood as a response to the problem of *quomodity*? Let us consider the factual aspect first. My answer then is 'yes and no'. There is nothing wrong with saying that models are similar to their targets, but it is not very telling to do so. The point is that the locution 'is similar to' functions as little more than a blank to be filled, since we must search for the appropriate replacement in each case. Or to put it the jargon introduced in the last chapter: similarity is an abstract concept that needs fitting out in every instance in which it is used and that has little interesting content of its own. A similarity claim is little more than an invitation to fill this blank by specifying relevant aspects of comparison, choosing an ordering in which these aspects enter and settling for standards of nearness. It is only once we have done all this that we know how model and target relate.[44]

Before briefly illustrating these claims, let me emphasise that I do not intend this criticism to establish that there is anything inherently wrong with similarity. I only want to point out that similarity *per se* does not provide us with a satisfactory answer to the (descriptive aspect of the) problem of *quomodity*. One might be under the impression that once we opt for similarity the problem is solved. This is wrong. What we need instead is a specification of scientifically relevant kinds of similarity, the contexts in which they are used, the claims they support, and so on. Before we have specifications of that sort at our disposal, we have not satisfactorily solved the problem of *quomodity*.

In saying that M resembles T one gives very little away. Everything resembles everything else in any number of ways. One can shelve one's books according to title, subject, year of first printing, publisher, colour, size, weight, price, or number of pages – this is a matter of convenience, taste and, in some cases, education. Similarity statements are vacuous without a specification of relevant respects. And this is by no means true only of humdrum examples like shelving books. Also in the context of science choosing the relevant respects in which two items are claimed to be similar is of crucial importance; and quite often this choice is not determined by the background against which a piece of research is carried out. Consider the following example from cutting edge physical chemistry. It has recently become possible to produce heavy elements, now known as transactinide elements, having atomic numbers between 104 and 108. In these elements, relativistic effects modify the structure of the outer orbitals in such a way that the chemical properties of a particular heavy element differ significantly from the chemical properties of other elements in the same group of the periodic table. In fact, its chemical properties are more like the properties of elements in another group of the periodic table than like the properties in the group it actually belongs to. So we are in the interesting situation that if atomic number is our concern, a heavy element, ${}_{108}Hs$ say, is similar to the elements in one group of the periodic table while it is similar to the elements in

[44] This seems to be in line with Giere's views on similarity. Though he does not put it that way, he is explicit about that fact that claims of similarity are vacuous without a specification of relevant respects and degrees (1988, 81).

another group when chemical properties are at stake. So it is not clear to which group it actually belongs, since chemical and physical criteria pull in different directions. This behaviour is very unlike that of 'lighter' elements for which the similarities in both respects coincide.

After having specified the relevant respects, similarity is still underdetermined. The problem is that it is not clear in what ordering the properties enter. One would expect that once we get down to the level of fairly specific, unidimensional features (or even simple properties) such as nucleon number, similarity is clear and straightforward. More to the point, one would assume that there is an ordering in which every simple feature unambiguously occupies one particular position and similarity is measured in terms of nearness in this ordering. This is false. Even when it comes to simple properties the assessment of degrees of similarity may vary considerably with context. What is pretty similar from one point of view may be far off from another. There are two aspects to this.

First, given a certain ordering, what is near and what isn't depends on the context and the aims we pursue. If we know the value of the constant of gravity with a precision of $\pm$ 0.001%, this is sufficiently accurate for an engineering task. But it is not accurate enough for some sophisticated experiment in particle physics, where a precision of at least $\pm$ 0.0000001% is required. So the value of the constant we have at hand is similar to the actual value in one context but not in the other. Or consider oscillations. In some contexts it is perfectly reasonable to describe the motion of a pendulum by a sine function, in others this is too coarse and we have to employ Jacobi elliptic functions.

Second, the assumption that there is one, and only one, correct ordering associated with each family of simple properties is wrong. As Goodman points out (1972, 445), even if we grant that the similarity of simple qualities can be measured by nearness of their position in an ordering, these qualities may be ordered in many different ways. Consider sounds. It seems straightforward to say that pitches are more alike if they differ by fewer vibrations per second. But this is only one way of looking at it. To a musician, middle C may be more like high C than like middle D. The same goes for colours. Physicists find it natural to order colours according to wave length. The closer the wave lengths of two colours, the more similar they are. But sensory similarity may not square with this ordering. To most people violet is more similar to red than to green, although red and violet belong to the opposite ends of the visual spectrum while green is somewhere in-between. Again others may take complementary colours to be similar. From this standpoint, red is more similar to green than to any other colour, and likewise for the pairs blue-orange, yellow-violet, etc. To sum up, even simple qualities like colour or pitch can be put in different orderings, when looked upon from different points of view.

The bottom line of this is that similarity is abstract. It is a dummy expression that needs to be explained in terms of something else every time it is employed. It is only after we have specified which features are relevant, chosen an ordering in which they enter and settled for standards of nearness that we know how model and target relate. So rather than similarity providing us with an explanation of the relationship between two items, this relationship, which has to be specified otherwise, provides a basis for a canon of similarity.

What about the normative aspect of the problem of *quomodity*? Is similarity a scientifically acceptable mode of representation or should we even require that all representations have to be of that sort. The answer to the first question is affirmative; if a model is similar to its target in a certain specified sense it certainly is a scientifically acceptable representation. But what about the second question? Given the abstractness of similarity I think that there is no way to tell. An answer depends on our canons of similarity. If we have liberal standards of similarity it may well be the case that all scientifically acceptable representations are of the similarity type; if our standards are tight, there may be models that do not qualify as such.

Part III

Re-Presenting Scientific Representation

Chapter 5

What Are Models?

1. Introduction

The bottom line of the discussion in Part II is that none of the currently available accounts provides us with a satisfactory answer to the basic problems of scientific representation. Nevertheless, the phenomenon of scientific representation is very real and our quest to understand how it works is legitimate. So we have to go back to the drawing board. This is what I do in this part of the thesis.

As a point of departure I take the result of the discussion in Chapter 3: a target system T possesses a particular structure S_T only relative to a certain description D. As I indicated, the expression 'relative to' stands in need of analysis (and the same goes for its synonyms 'with respect to' or 'against the background of'). What exactly does it mean for a target system to have a structure relative to a description? This is a question a theory of representation has to answer; but over and above being important in its own right, it serves as a springboard to a discussion of the ontological puzzle (as introduced in Chapter 1). To come to terms with these two problems is the aim of this chapter.

The answer I will suggest is simple: models are objects, either imagined or physical. The former claim is a consequence of the view on how objects have structures that I developed in Chapter 3. For this reason, the bulk of this chapter is devoted to models as imagined objects. I first introduce this point of view (Section 2) and then defend it against the criticism that this is an unnecessarily inflated ontology and that models are, after all, nothing but structures plus descriptions (Section 3). Then I present a second argument for the conclusion that models are imagined objects, which is independent of the first one and which should satisfy those who were unconvinced by my arguments in the previous sections (Section 4). I then discuss another objection, which has it that my account is incomplete because there are things – material objects and equations – that qualify as models and that do not fall within the group of imagined entities (Section 5). I argue that the objection has a point as regards material objects but fails when it comes to equations. For this reason, we also have to include material objects in the class of things that can be models. This is no problem, however, because material and imagined models can be covered by the same theory of representation. I close this chapter by pointing out in Section 6 that the views developed here nicely square with the similarity view of representation.

2. Models as Imagined Objects

The conclusion of Chapter 3 is that a target system has a certain structure *S* only relative to a more concrete description *D*. Or more specifically, assuming that *D* is the more concrete 'fitting out' of *S* in some particular situation, it is true that the target *T* has structure *S* only if *D* is true of *T*. So far so good. The problem with this is that most descriptions we use in connection with scientific models are not true. It is a commonplace that most models involve idealisations or approximations of some sort or another; the examples presented in Chapter 3 illustrate this point. The consequences of this are severe. If the descriptions used to 'ground' the structures involve idealisations, say, then they are not true, and neither are the structural claims they fit out. Taken at face value, it is false that planets are spheres or that generations reproduce at constant rate and therefore the structural claims based on these assumptions are false as well. So the targets actually do not possess the structures we took them to possess.

This is an unwelcome consequence. A theory that rules out all but perfectly accurate representations is unacceptable. But how can we evade this conclusion? My suggestion is the following. Instead of taking the description to be a description of the target system itself – taken as which it is false – we should take it to describe an imagined object of which it is true by assumption. Then we posit that this imagined object represents the target system. For this reason this imagined entity *is* the model. Given this, we can replace the notion that a target system *T* possesses structure S_T

relative to a certain description D by the (more liberal) notion that a target system T possesses structure S_T *as represented in model M.*

As a simple example consider the usual rendering of the solar system describing planets as spherical mass distributions that only interact with the Sun, etc. On the suggested reading, the way to look upon that is to say that this description canvasses an imagined entity, namely one consisting of ten perfect spheres, nine of which orbit around the tenth etc. This imagined entity is a model representing the solar system. Because all the claims the description makes are true of that model by assumption, the structural claims are true of it as well. Hence the imagined entity has the structure for which the description is a fitting out. The target system itself then has this structure in a derivative way, namely *as represented* in the model.

In sum, there are four claims involved in this suggestion. First, false descriptions describe imagined entities. Second, this imagined entity represents the target system. Third, this imagined entity is the model. Fourth, a target system T possesses structure S_T *as represented in model M.*

These claims need to be spelled out and argued for. The first, the third and the fourth claim are only seemingly problematic and I will be dealing with them in this section. The second claim is the real challenge. How does an imagined entity represent a target system? The next two chapters are dedicated to a discussion of this problem and therefore I will not say more about it now.

False descriptions describe imagined entities. That false descriptions of the sort involved in scientific modelling can be understood as descriptions of an imagined entity seems to be obvious and unproblematic. We all understand the idealised descriptions of the solar system and know what an entity of which these claims are true would be like. And similarly for a box full of billiard balls, beads connected by springs, and so on. These models, however, are in the scientist's mind rather than in a laboratory; they do not usually have to be actually constructed and experimented upon to perform their representational function. It is sufficient that they exist in a scientist's imagination. For this reason it seems to make sense to refer to models of this sort as 'imagined models'.

Two qualifications are needed, however. First, there are time-honoured worries about the ontological status of 'things in the mind' and to come up with a tenable ontology of mental objects is a veritable quandary in the philosophy of mind. Though interesting in its own right, this problem need not worry us in the context at hand. When it comes to scientific modelling, mental objects are derivative in the sense that what really matters is the model as an object, not the fact that it is imagined. What Maxwell was dealing with was billiard balls *tout court*, not mental images of billiard balls. Referring to these as an 'imagined model' should only indicate that Maxwell never had a vessel full of billiard balls in his laboratory. He thought about billiard balls, that's all. But it would be a mistake to think that it was his *mental image* of the billiard balls that represented the gas. It is the *billiard balls themselves* that do the representational job. So by taking models to be imagined entities we are not committed to absurd claims of the sort that there are items in the mind that possess spatial dimension and other physical properties and that can be rotated or moved. It is the objects themselves that possess these properties, not their mental images, and it is these objects that serve as scientific representations, not their mental counterparts. Moreover, we have firm pre-theoretic intuitions about imagination that we can build on. I take it to be uncontroversial *that* we can imagine certain things – a vessel full of billiard balls, for instance. Whatever difficulties may come up when we try to explain *how* this is possible, the phenomenon of imagination is very real, and it is this pre-theoretic understanding of imagining something that I am building on here. Nothing over and above this is needed for the purpose of the present discussion.

Second, I should be careful to point out that the class of imagined entities does not coincide with the class of fictional entities; rather the latter may be considered a subclass of the former.[45] By fictional entities one typically means entities originating in myths, fairy tails, or novels such as unicorns, Madame Bovary, or Mickey Mouse. They are entities that lack existence in the physical world but which nevertheless can be the subject of meaningful discourse. Some scientific models (or some of the entities figuring in them) belong to this group of entities as well: frictionless planes, point masses, pendulums with mass-less strings, spherical planets, the *homo oeconomicus*, etc. These few examples indicate that fictions play a rather important role in scientific reasoning. It is not clear, however, in what way they do so. Fictional

[45] What I have in mind at this point is how we have access to these entities. A Platonist would hold that fictional entities exist irrespective of whether someone imagines them or not, while a constructivist would deny that. Both would agree, however, that it is via our imagination that these entities become present to us.

entities are notoriously beset with a host of difficulties. How to draw the distinction between real and fictional entities? Can we grant them some ontological status or do we have to eliminate them in rational discourse? If the former, what is their ontological status and what kind of existence (or non-existence) do they have?

These problems do not have to be solved now. What matters for the time being is the fact that fictional entities fall under the category of imagined entities, but not vice versa. The first part of the claim is trivial. We imagine frictionless planes or point masses in the same way in which we imagine billiard balls. However, not all imagined entities need to be fictional. We can imagine entities that actually do exist (billiard balls) or we can imagine certain things that do not yet exist but could be constructed, at least in principle (some particular experimental set-up, for instance). In other words, there is no presupposition that imagined models *cannot* exist in the actual world. They may not exist (yet) for some reason or another but nothing in the notion of an imagined model requires that this be so.

Models are imagined entities. This claim comes for free. A model is the thing that represents a target system. If we have come to the conclusion that it is an imagined entity that performs this function, then this imagined entity is the model.

A target system T possesses structure S_T as represented in model M. This is a direct consequence of what has been said so far.

3. The Descriptivist Objection

Descriptivism has it that the whole talk about imagined objects is totally misguided and puts forward the view that models are, after all, nothing but descriptions. On this view, what scientists display in scientific papers and textbooks when they present a model are descriptions. These descriptions may be different in character, use different languages, and vary in accuracy, but – and this is the salient point – they are descriptions. Nowhere do we need dubious entities like imagined objects. Models are descriptions of some system, no more and no less. The introduction of an additional non-descriptive layer between our descriptive language and the world, containing imagined objects, is unjustified and leads to an unnecessarily inflated ontology.

These objects, so the objection continues, are not doing any work and we should dispose of them or, even better, not introduce them in the first place.

In this section I argue that this objection fails because whenever one tries to render the view that models are descriptions precise, one cannot evade appeal to imagined objects at one point or another.

On the view at stake, models are descriptions of a system. But are they really? To reiterate the point made above, in many cases the descriptions used in science are simply false. Planets are not spherical, surfaces are not frictionless, a polymer is not a collection of beads connected by springs, and so on. If we take models to be descriptions, it is not clear what they are descriptions of. At any rate, they do not seem to be descriptions of the target system.[46]

The descriptivist may now counter that this objection rests on the unwarranted premise that descriptions have to be literal. There are many different ways of describing something and coming up with a literal description is only one of them. Then, there is no reason to require that scientific descriptions have to be literal descriptions. All we need, even in a scientific context, is a description of some sort, literal or not, that presents the system to us in a cognitively significant way.

This may amount to three things. First, one may take non-literal descriptions to be approximate descriptions of some sort. Second, one could understand them as metaphors. Third, one could consider them to be a set of assumptions about a system. I now discuss each of these suggestions and conclude that they are of no help when we aim at getting rid of imagined objects.

Consider approximate descriptions first. One obvious way of circumventing the introduction of imagined objects seems to be to reply that although the descriptions involved are not literally true, they are approximatively true, and that this is good enough. This suggestion suffers from the problem that it invokes a highly problematic notion, namely approximate truth. Much ink has been spilled on this issue but no widely accepted account has emerged from the debates. The descriptivist may now reply that this is asking for too much. What we need in order to deal with the problem at hand are some firm intuitions about what it means for a description of

[46] This has also been pointed out by Giere (1999, 122-3).

the kind needed in scientific modelling to be approximately true rather than a general account of approximate truth. Let's grant this point and see where it takes us.

How would we make a judgement about the approximate truth of a description *D*? The most natural thing to do in the given context, I take it, would be to make a comparison between the thing literally described, *A*, and the thing of which the description is supposed to be approximately true, *B*. Then one can say that *D* is approximately true of *B* if *B* is close to *A* on some previously accepted standards of proximity. For instance, we would consider the above description of the planetary system to approximately true if we come to the conclusion that, on some standards of proximity, the shape of real planets is close to a perfect sphere.

If one grants that this (or something along these lines) is correct, then the descriptivist has given the game away. Approximate truth rests on the relation the target bears to another object, which we have to construe as an imagined object because it does not usually exist in the physical world. Hence we have gotten back the imagined objects we wanted to get rid of. Now have a closer look at the nature of the relation between the two objects. What makes *D* approximately true of *B* is that *A* is close to *B* on some standards of proximity. This amounts to saying that *A*, on these standards, could stand in for *B*, that it portrays certain aspects of *B* sufficiently accurately, or that it manifests certain features of *B* to some specified degree. And this in turn amounts to saying that *B* represents certain features of *B* in some way. So we are back to the picture I sketched above. The description literally describes an imagined object and this imagined object represents the target in some way. For this reason, I think it is appropriate to reserve the term 'model' for the object literally described by the locution and to refer to the locution itself as a description of the model.

Next, consider metaphors. Not all of these descriptions can be understood as approximate descriptions; some have clearly metaphorical character. Can these be understood descriptively? I now argue that when understood in the descriptivist vein they suffer from the same difficulty as approximate descriptions: when accounting for when the metaphor is true we need imagined objects. To see how these enter the scene, reflect on how we understand and subsequently put to use metaphors. Within the context of science,[47] we understand how to use a metaphor if we are able to specify (at least tentatively) what the relationship is between the thing literally described by the locution and the object it is applied to metaphorically. As an example, consider a famous metaphor from seventeenth century cosmology, which has it that the universe is a clockwork. We understand the metaphor once we realise that certain properties of the clockwork – evolving according to a fixed and invariable plan, being designed by an intelligent agent to do exactly that, leaving no room for deviation, etc. – are also believed (or were believed) to be properties of the universe.[48] Though false as a literal description, the locution makes sense as a metaphor because we can compare aspects of a clockwork to the universe. And the same is true of beads on springs, billiard balls, etc. We can make sense of the locution 'the universe is clockwork' but not of 'the universe is a carburettor' because we claim that the universe has some interesting aspects in common with clockworks but not with carburettors. The richer and more sophisticated the specification of this relationship is, the more telling the metaphor becomes. The details of this specification do not matter in the context at hand, nor does the epistemic question of

[47] I do not claim any of this to be valid outside the realm of science.

[48] In making this point I am using the classical theory of metaphors, stated in Aristotle's Rhetoric and now commonly referred to as 'comparison theory'. On this view, every metaphor involves a comparison between two things, one of which is designated literally by a locution and another one which is designated metaphorically by the same locution. But the point I am making does not depend on the endorsement of this particular theory of metaphor; it can equally be made from different points of view. Let me briefly indicate how. According to another influential view on metaphors, the interaction theory, every metaphor involves a semantic interaction between a literal element in a sentence and a metaphorical element. The problem with this theory is that 'interaction' in this context is itself a metaphor, and hence the theory does not account for how metaphors work. For this reason, more recent versions of the theory explain interaction as the placement of the object literally denoted within the conceptual system associated with the metaphorical term. Returning to the above example, this amounts to associating the universe with attributes of clocks. And this is, at least in the cases at stake, just a different way of saying that we ascribe properties of the clockwork to the universe, which is the point that I am making. According to the third influential view on metaphors, the speech act theory, it is not words or sentences per se but the use we make of them in a specific situation that is metaphorical. The analysis of the particular locution and its use within speech act theory is a convoluted matter, one that I cannot get into here. But I take it that whatever one wants to say about how an expression like 'the universe is a clockwork' functions metaphorically, at some point one has to engage in comparisons between the object literally denoted and the one referred to metaphorically.

how we come to know. The salient point is that we make sense of a literally false locution by pointing out some relationship between what it literally describes and what it is metaphorically applied to.

Now we are in the same situation as above when explaining how approximate descriptions work. We posit an entity over and above the target system and start specifying in what relation this entity stands to the target system. And as in the case of the approximate description we have to say the metaphor is true if this 'secondary entity' shares some interesting properties with the target (at least to some degree); and this means that the secondary entity represents (certain aspects of) the target. Again, we are back to picture I suggested in Section 1.

The third suggestion is to understand a model as a set of assumptions about a system. In this vein Peter Achinstein suggests that 'when scientists speak of a model of *X* they are not referring to some object or system *Y* distinct from *X*, but to a set of assumptions about *X*' (1968, 212). This is of no help in the present context. If the 'assumptions' the suggestion refers to are true, then we do have a plain description and there is no problem to begin with. But referring to a claim as an 'assumption' characteristically implies that one takes it to be literally false. But then we are back where we started; and unlike the other two suggestions (approximate descriptions and metaphors), construing false claims as assumptions does not provide us with any clue as to how false claims should be related to the actual target system. At best we can understand it as an invitation to consider a system that possesses a certain set of properties of which we believe, given our best knowledge, that they bear some interesting relationship to the properties that the target system in fact possesses. Then it is reasonable to understand the former as a model of the latter. But now we are in the same situation as in the previous cases. What we are really doing is comparing two objects, one real and the other imagined, and it seems the most natural choice to say that the latter is a model of the former.

From this I conclude that one cannot do without imagined models. The descriptivist may now grant this point but insist that not all cases work in this way. He may object that the discussion so far has been tailored towards the needs of highly unrealistic models such as the beads connected by springs and that, even if we assume that it gets things right for this kind of models, it does not cover cases of more realistic scientific representation. At least in cases in which the description is more or less accurate, we can dispense with the extra layer of imagined objects and take the description itself to be the model.

Admittedly, this objection is more difficult to counter than the ones I have been discussing so far. My reasons for thinking that my analysis is correct even in cases where the description is realistic are the following. To begin with, I doubt that there are many descriptions in science, if any, that are realistic enough for us to believe that the description describes reality itself rather than a 'secondary entity' of sorts. The paradigm of a successful and realistic model is the one of the solar system. But still, the idealisations made are considerable (no gravitational interaction between planets, all planets are spherical, etc.) and for this reason there is no way a description of the model could equally well serve as a true description of reality itself. And this is equally true of most, if not all, scientific models. Therefore, the objection, even if correct, does not seem relevant for scientific practice.

One may counter, legitimately I think, that this is not a satisfactory argument because what we are aiming at here is a conceptual analysis of scientific representation and not an inventory of factually available models. There is a second difficulty for the descriptivist, however, which cuts deeper than the factual unavailability of realistic descriptions. Where do we draw the line between descriptions that are realistic enough to be considered descriptions of reality itself and ones that are not? If we grant that in the case of unrealistic descriptions the analysis that I have presented above is correct, we have to specify at which point the transition from descriptions that describe models to ones that face reality directly takes place. This, I think, cannot be accomplished. There is no way to specify how realistic is realistic enough. Every choice we may make is somehow arbitrary. Some may now want to say the 'transition condition' is approximate truth; that is, if the description is approximately true, then it can be considered a description of reality itself rather than of a model. However, given that no tenable account of approximate truth is available as yet, this suggestion seems to be a non-starter. But what alternative do we have? I can't think of any.

For these reasons it seems best to adopt the analysis I have proposed uniformly for all models, regardless of how realistic they are.

4. A Faster Route to Imagined Models

I will now present a second and more direct argument for the conclusion that models are fictional objects, which is independent from the argument previously presented. The present argument may convince those who feel that what I have presented so far is too idiosyncratic to my own way of looking at things and that it makes too many assumptions (about approximate truth, metaphors, etc.) in order to be conclusive.

To begin with, let us have a closer look at the descriptions involved. When recapitulating the examples proffered in Chapter 3 to illustrate the point that structures rest on descriptions, we immediately realise that these descriptions are plain and straightforward specifications of what the basic constituents of the system are, of what properties they have and of how they interact. Moreover, as the same set of examples shows, these descriptions for the most part involve simplifying, approximating or idealising assumptions. We describe atoms as point particles, planets as spheres, generations as discrete unities, and so on. When put in this way, it is obvious that what we are faced with when dealing with descriptions of this kind are models of the target in the straightforward scientific sense of the term. Putting forward assumptions of the above kind is what scientists do when they present what they would call a model of the system they are interested in.[49] To a physicist, a model of the solar system is the above-mentioned collection of spinning spheres and to a biologist a model of a population is a sequence of discrete generations. At the most basic level, what scientists mean by a model of a target is a set of assumptions about what entities the target consists of, what properties it has, what interactions take place and what claims it satisfies.[50]

So the concrete descriptions fitting out a structural description are what scientists commonly call a model. What is the relevance of this observation to the ontological

[49] Compare the following definition of a model taken from a standard physics textbook: 'In physics a model is a simplified version of a physical system that would be too complicated to analyse in full detail.' (Young and Freedman 2000, 3).

[50] This conception of models is not foreign to the philosophical literature either. Achinstein's theoretical models (1968, 209), Cartwright's prepared descriptions (1983, 133-4), Hesse's analogical models (1963, Ch.2) and Suppes' physical models (1970, Ch. 2 p. 9) are closely related – not in concrete detail but in spirit – to the scientist's use of the term 'model'.

puzzle? Does it tell us what models are? Not as it stands. Two readings are possible. On the one hand, one could take the above literally and claim that the model really is the description itself. On the other hand, one could interpret it as saying that the model is the thing described by the description rather than the description itself. In the remainder of this section I argue that the first reading is untenable and that we should adopt the second. This drives my point home, since when we look at what these descriptions are about, we realise that they describe fictional entities.

The claim that models are descriptions is flawed for two reasons. First, if we identify a model with its description, then each new description yields a new model. But we are all familiar with the fact that the same thing can be formulated in a number of different ways. For instance, one can translate a description into other languages (formal or natural), but one would not say that one thus obtains a different model. When, for instance, we are faced with a French translation of the above description of the solar system, the words and sentences we read are different than the ones in the English original. Hence it is a different description. Of course, both descriptions describe the same thing (at least if it is a good translation), but that is a different matter. The descriptions *as descriptions* are different and this renders the notion that models are descriptions highly implausible. We don't get a different model every time we translate our description into another language. This naturally leads to the conclusion that the model is what the description describes and not the description itself.

Second, models have different properties than descriptions. We say that the model of the solar system consists of spheres orbiting around a big mass, that the population in the model is isolated from its environment, or that the force connecting two beads in a chain is harmonic. But these statements are sheer non-sense when we take models to be descriptions. The description itself is not spherical or isolated from the environment and it has no harmonic forces in it. It is things described by the description that possess these properties, not the description itself. On the other hand, descriptions have properties models do not have. A description can be written in English, consist of 517 words, be printed in red ink, and so on. None of this makes any sense when said about a model. For this reason, models and descriptions are not the same.

The conclusion is that descriptions and models are distinct. The model is what a description describes, not the description itself. This leaves us with the question of what it is that descriptions describe. But this question is not new; it has been discussed above and the answer given remains valid in the present context. The descriptions used in science describe imagined objects. This is because these descriptions are characteristically false when understood as literal descriptions of the target system and for this reason it is best to construe them as describing imagined objects in the sense specified in Section 2.

5. Other Kinds of Models

Another objection to my view that models are imagined objects is that it is incomplete. When we look at what passes as a model in scientific discourse, we find at least two more kinds of things over and above imagined objects that are habitually referred to as 'models': material objects and equations. There is no reason, so the objection goes, to deny them the status of a model. For this reason my account is too narrow.

My reply to this is the following. As to material objects, the objection has a point. It is true that some important models are material objects and that I have not said anything about them as yet. This sin of omission is easy to fix, however, because material and imagined objects are close enough to each other to carry over most of what has been said and of what will be said about imagined objects to material objects without much ado. As far as equations are concerned, things are less straightforward. Although equations are habitually referred to as 'models' and there seems to be no harm in doing so for practical purposes, from a philosophical point of view this is mistaken. Equations are syntactic items and as such they face objections similar to the ones marshalled against descriptivism in the previous sections. Let me discuss these two points in turn.

Material Models

The most straightforward kind of models is material objects. For want of a better term I refer to models of this kind as 'material models', where 'material' is understood in the broadest sense possible. The class of material models comprises anything that is a physical entity and that serves as a scientific representation of something else. Among the members of this class we find stock examples like scale models of bridges or planes and analogue models like electric circuit models of neural systems or pipe models of an economy. But also more cutting edge cases, especially from the life sciences, belong to this category. It is common practice in biological and medical research to use some organisms to study certain aspects of other organisms. Snails are investigated to understand aspects of the nervous system of primates; patterns of the development of cells in mammals are studied in worms (I discuss this case in the Chapter 7); and mice are used to test drugs that are designed for humans. From a semantic point of view this is to say that snails, worms and mice serve as models of primates, mammals and humans.

The crucial thing to realise is that from a representational point of view, there is no difference between imagined and physical objects (and for this reason I sometimes use 'object-model' as an umbrella term covering both material and imagined models). That is, it is irrelevant to the question of how representation takes place whether the model is a real physical entity or whether we 'merely' imagine it. I observed above that when we deal with imagined entities it is still the entities themselves that do the representational job; it is the billiard balls themselves that represent the gas, not our mental image of them. It is the objects themselves that possess properties, not their mental images, and it is these objects with these properties that serve as scientific representations, not their mental counterparts. But if it is the objects themselves that do the representational job, it cannot matter whether the object is real or imagined. Whether we have a couple of beads connected by springs in a laboratory or whether we postulate such an entity only theoretically does not make any difference to how it represents a polymer (as long as the imagined and the real entity have the same properties). And similarly in the case of the solar system. It is irrelevant to the question of how spinning spheres scientifically represent planets whether there actually are such spheres or not. We may find a realisation of the model in the science museum, but this merely serves pedagogical purposes and has no bearing on either of the semantic aspects (enigma and *quomodity*) of the model. The physical thing in the museum and its imagined counterpart represent the solar system in exactly the same way. In short, from a semantic perspective material and imagined models are equivalent.[51]

This, of course, still leaves us with the question of how objects represent – which I address in the following two chapters – but we now know that whatever answer we will come up with, it will be valid for objects irrespective of whether they are real or imagined.

A cautionary remark should be added here. The fact that the difference between material and imagined models does not matter to their semantics does not imply that this difference does not matter in other respects. Trivially, it is of fundamental importance to ontology. But it also has important consequences for the epistemology of models. Learning from material entities we can experiment on takes place in a way that is very different from gaining knowledge by dint of reasoning about an imagined entity. But all these differences notwithstanding, from a *semantic* point of view the two work in exactly the same way and this is what matters for now.

Equations

Another group of things that are habitually referred to as 'models' is mathematical equations. In particular in economics this is the standard usage of the term. What is referred to as 'the model' in a paper or a textbook is usually a set of equations. The Black-Scholes model of the stock market or the Mundell-Fleming model of an open economy are cases in point.

As long as this is taken to be no more than working scientists' jargon, there is no harm in referring to an equation (or a set thereof) as a 'model'. But taking this habit as guide to a philosophical analysis of models is seriously misleading. Equations are syntactic items, they are descriptions of some sort, and as such they are open to the

[51] It may well be – and normally is – the case that we cannot produce a material equivalent of a certain imagined model, a frictionless plane say. Therefore imagined and material models are normally different. But, and this is the salient point, they are not different because they have a different *semantics*. They are different because the objects that serve as models are different. If we were able to produce a frictionless Euclidean plane in the laboratory, it would represent the real slope for the same reasons and in the same way as its imagined counterpart.

same kind of criticism as descriptivism in Section 3. First, one can describe the same situation using different co-ordinates and as a result obtain different equations. As an example consider the motion of a three-dimensional oscillator. The equation we obtain when we write down the equation of motion in Cartesian co-ordinates is different from what we find when we write it down in spherical co-ordinates. Or one can use different formalisms altogether to describe the same situation. For instance, one can write down the equation of a simple linear electromagnetic wave in standard vector calculus, or one can use the formalism of tensor calculus. In all these cases the equations look very different although they describe the same situation or object. Second, the model has properties different from the equation. The above oscillator is three-dimensional; but the equation describing its motion is not. The model is continuous; but the equation is not. On the other hand, an equation may be inhomogenous, the system it describes is not. An equation may contain a diagonal matrix, the resonance problem it describes does not. And so on.

For these reasons I deny that equations are models, although this contradicts the use of the term in scientific practice. As in Section 3, the conclusion is that a model is what an equation describes rather than the equation itself. But at this point the parallels between these two cases break down. In contrast to the descriptions discussed in section 3, which can easily be understood as describing imagined entities, it is not clear at all what equations describe. Here we are getting into deep waters. The question of what equations 'are about' is no less than the time-honoured question of what the subject matter of mathematics is; and the question of how equations can be used within an empirical science is the puzzle of how mathematics applies to the physical world. These questions deserve serious consideration and I come back to them in detail in Chapter 8, where I propose an answer to both problems and explain how these answers square with my views on modelling developed in the other chapters of this thesis.

Summary

In sum, there are two types of things we can legitimately call 'models': material objects (e.g. the model of a car we use in experiments in the wind tunnel) and imagined objects. The latter further divide into fictional models (frictionless planes, mass less strings) and models comprising entities that do actually exist in our world, but are only imagined by us for the purpose of the model (beads connected by springs). For want of a better term I call such models 'non-fictional'. So we obtain the following picture:

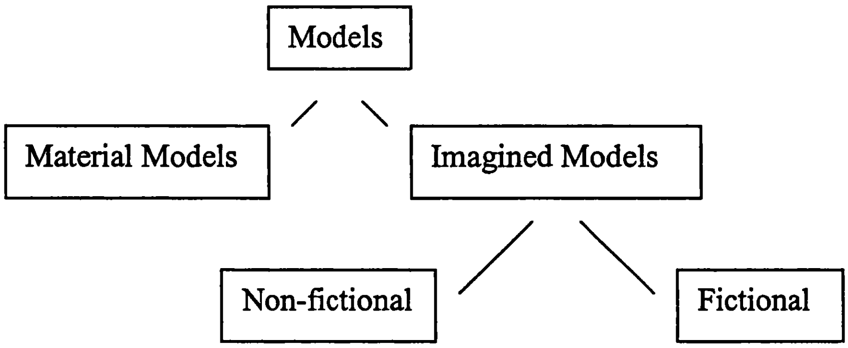

Fig. 4: The ontology of models.

6. A Brief Remark on Models and Similarity

I would like to end this chapter with a remark about the similarity view of scientific models. So far the discussion has been geared towards the structuralist version of models. How does what I say in this chapter bear on and square with the similarity view?

As I see it, the views developed in this chapter are perfectly in line with the similarity view on what models are. Objects more naturally enter into similarity relations with target systems than structures and for this reason a similarity theorist should naturally be drawn towards thinking about models as objects. So it comes as no surprise that Giere repeatedly paraphrased models as 'objects' (2002), or 'abstract entities' (1988, Ch. 3). Structures no doubt play an important role in science but, as far as I am aware, not much has been said within the similarity paradigm about how this squares with a view that takes models to be objects. In this respect, the discussion in this chapter can be understood as an improvement on the similarity notion of models because it clarifies the relation between concrete objects and structures. This is important because, as I will detail in Chapter 8, this is needed in order to account for the use of mathematics in science.

Chapter 6

On Scientific Aboutness

> 'Art is no longer a purely visual sensation that we record, a photograph of nature, as sophisticated as possible. On the contrary, it is a creation of our spirit which nature provokes. [...] Art, rather than a copy, becomes the subjective transformation of nature.' (Maurice Denis 1909, 48-49.)

1. Introduction

In the last chapter I argued that models are objects, either imagined or physical, which represent their target systems. This takes us back to the enigma of depiction as formulated in Chapter 1: due to what does a model represent a target system? This is the problem I deal with in this chapter and the next one. The division of labour between these two chapters is as follows. The present chapter has preliminary character in that it outlines some of the most basic features of scientific representation. In doing so it points in the direction in which we have to look for a constructive account and establishes conditions of adequacy: no account that runs counter to what is said about scientific representation in this chapter is acceptable. In doing so this chapter paves the ground for the formulation of a positive account in the next chapter.

To give an outline of the most basic features of scientific representation involves two things. On the one hand, we have to bring to the fore some general facts about scientific representation. Some of them may seem obvious, others less so. But in either case it is important to make them explicit right at the beginning in order to evade pitfalls later on. On the other hand, it is imperative to diffuse some misconceptions about scientific representation before they can seriously mislead us.

A word of warning seems in place at this point. Representation is a broad notion. In fact, just about anything that can be semantically evaluated can be called a representation: words, sentences, paintings, photographs, sculptures, diagrams, graphs, equations, charts, road signs, maps, signposts, scale models, gestures, facial expressions, acoustic signals, etc. There seems to be almost no limit to the range of representational systems human beings are able to devise and use. But these are not all of the same kind and there is no reason to assume that they all work in the same way. For this reason, a great deal of the discussion in this chapter is concerned with setting off scientific representation from other forms of representation. This is insightful because we learn about scientific representation by learning what it is not. This, however, is only aimed at identifying something like 'symptoms of the scientific' and does not provide us with a set of necessary and sufficient conditions for a representation to be scientific. The characteristics I describe in this chapter are at best necessary for a representation to be scientific, and certainly far form sufficient. Other forms of representation equally meet these requirements. The characteristics outlined in this chapter define a rather broad class of representations of which scientific ones form a subclass. A discussion of what – if anything at all – marks this distinction must be left for later.

2. 'Naturalness' and the Acquisition of Knowledge

The main point I am getting at in this section is that scientific representations belong to a class of representations which work in a way that the properties they themselves, as objects, possess are crucial to the performance of their representational function. I reach this conclusion by first observing that scientific representations function cognitively and then asking how they can do this. In discussing this question I first dismiss the suggestion that scientific representations have a certain kind of 'naturalness' and then discuss how models differ from lexicographic representations. These considerations finally lead me to the conclusion just mentioned.

Different representations serve different purposes. Some are devised to please the eye; others serve the purpose of communication; and again others are used as objects of religious devotion or means of ideological identification. In contrast to these, scientific representations function cognitively. Their purpose is to instruct us about the things they represent. They do not merely stand for something beyond themselves; they present things to us in a way in which we come to understand them and acquire knowledge about them. Scientific representations are, as Mary Morgan and Margaret Morrison put it, investigative tools (1999, 11). We study a representation and thereby discover features of the thing it stands for (its target system); that is, they allow for what Chris Swoyer calls 'surrogative reasoning' (1991, 449). Or to put it another way, we can, as it were, 'see through' scientific representations to the target behind them.

How is this 'seeing through' possible? How can a representation function cognitively? A first reaction to this question might be to reply that scientific representations are somewhat 'natural' (for instance in a way similar to how figurative paintings are said to be natural). That is to say that they are icons or mirror images of sorts in that they present things to us (roughly) 'as they are'. Although I am not aware of anybody explicitly putting forward this view, at least not in the crude form in which I am presenting it here, it is worth some consideration. This is first because we learn a great deal about the character of scientific representation by refuting it; and second because it is the extreme version of a family of related views that, though rarely if ever explicitly articulated, seem to loom in one way or another in the back of the minds of at least some proponents of the semantic view.[52]

The intuition backing this view is the observation that scientific representations seem to be less conventional than other signs. Unlike lexicographic representations (linguistic representations, for instance), which can be chosen more or less arbitrarily and which receive their significance through arbitrary stipulation and use, scientific representations seem to be – to a great extent at least – determined by the features of the target. We do not seem to have the same freedom in choosing models as we do in choosing verbal symbols, say. What lexicographic representations we choose to represent something is a matter of consensus and we can replace one by another one if we please without thereby changing the way in which things are represented and the way in which we learn about them. The cognitive import of lexicographic symbols entirely depends on their conventionally established semantics and the symbol *as an object* is cognitively insignificant. This becomes clear from the fact that we can translate a sentence into many different languages while still saying the

[52] French, for instance, explicitly refers to figurative pictorial representation to vindicate his isomorphism view of scientific representation (2002).

same thing. The sentence *as a physical entity*, i.e. the sentence as a particular string of signs on paper, does not contribute to how a certain matter of fact is portrayed, it is the conventions associated with it that do.[53] We cannot, however, arbitrarily replace one model by another. Something over and above sheer stipulation seems to be involved in the way in which a model represents its target. And this extra element, so the suggestion goes, is a kind of 'naturalness', which ties models and their targets much closer together than words and their referents. A model must, as it were, naturally lead the scientist to the target and inform him about it. Even if one grants that this is not possible without convention of sorts, there seem to be clear limits to the role of convention. Convention alone cannot make a model. And it is for this reason, so the suggestion concludes, that models can inform us about the properties of the target. We do not learn anything about an object by studying conventional signs by which it is denoted. This is because there is nothing but bare stipulation to them. Models are different, however, since they have the component of 'naturalness' to them. It is by looking at these natural features that we can learn about the target.

This view, plausible as it may seem at first glance, is untenable for at least two reasons. First, it is not clear what naturalness is. Where do we draw the line between the conventional and the natural? What is it for something to be a mirror image of something else? As far as I am aware, there are no satisfactory answers to this question available. Similarity and isomorphism could be understood as attempts to draw a line between the natural and the conventional. But even if we were prepared to grant that they get the natural-conventional distinction right, they do so, as I have argued in the previous chapters, in a way that is uncongenial to the needs of a theory of scientific representation.

The second problem is that even if, for the sake of the argument, we assume that a satisfactory account of naturalness were available, we will have difficulty finding models that would qualify as natural. Wherever we come to draw the line between natural and conventional, many scientific models clearly fall on the conventional side. Just consider the set of models I started off with. Polymers are complex chemical compounds and there is very little natural in representing them as beads

[53] In passing I should mention that this has led some philosophers of language to distinguish between sentences and propositions, where propositions are construed as abstract entities distinct from the sentences expressing them.

connected by springs; molecules, in particular if they consist of more than one atom, have relatively little to do with billiard balls; and the Lorenz model of the dynamics of the atmosphere is based on so many simplifying assumptions that we would hardly consider it to be natural. Moreover, in many cases we just don't know whether the model could even be considered natural or not. Atoms, quarks, and the universe as a whole are not accessible to direct experience and we may therefore not be able to tell how natural a model of one of these entities is. Finally, even in cases in which we would grant, at least intuitively, that there is a certain naturalness to a model, we realise upon closer examination that some conventions are at work as well. There is nothing in the nature of planets that renders the choice of spheres rather than ellipsoids, say, more natural; and the decision to neglect the gravitational interaction between the planets is motivated by mathematical considerations rather than by the physics governing these objects.

For these reasons, models cannot be considered natural representations of their targets. However, one cannot resist the impression that all these arguments are somehow contrived and that there is something to the view that models are natural representations. Even if we admit that conventions play some role in scientific modelling there seems to be a considerable difference between sentences and models and this difference seems to have something to do with the 'naturalness' of models. This raises the question of whether there is something to this intuition or whether it really is nothing but a relic of a misguided view on models.

The grain of truth in this view, I think, is the following. What the natural-conventional divide really tries to capture is the distinction between representations whose internal constitution matters to their functioning as representations and ones for which this is not the case. The properties of a word (consisting of so and so many letters and syllables, occupying this or that position in a dictionary, etc.) do not matter to its functioning as a word – and likewise for any other lexicographical sign. We can replace one word by another according to will. This is possible because the properties of a word *as an object* do not matter to its semantic function; that is, what a word stands for is in no way dependent on the features it possesses as an object. For this reason, words by themselves are uninformative. We can twist and turn the word 'atom' as long as we wish and we will not learn anything about atoms.

This is very different with models. The properties a model has do matter. Unlike with lexicographic representations, we actually investigate the model itself to find out about its target. Models are entities with an internal configuration or set-up and it is by exploring this internal configuration that we learn about the target. The model of the planetary system consists of spherical spinning tops with spherical mass distribution that interact gravitationally and it is by exploring these features that we learn about the planetary system. If we change one of these features the way in which the model represents the target changes as well, and with it the means by which it instructs us about the target. For this reason, the properties a model possesses play a crucial role in the performance of its representational function.

For this reason models may seem to represent in a more 'natural' way than words. This, however, is a faulty conclusion. The fact that the properties models possess matter to how they represent their targets does not mean that they do so in a more natural way than words. But it brings to our attention the matter of fact that models represent in a way that is very different from how words represent. What this way is will be the subject of the next chapter. Before discussing this problem, a few more preliminary questions need to be settled.

3. The Intentional Character of Scientific Representation

So far I have argued that models are objects, either imagined or material. But not every object is a scientific representation. So what marks the difference between objects that are scientific representations and ones that are not? (The points I am making in this section are valid irrespective of whether we consider imagined or material objects. For this reason I talk indiscriminately about 'objects' and drop the qualifications 'imagined' or 'material'). Two types of answers can be given to this question, a naturalist and a non-naturalist one. On the naturalist view, objects that are scientific representations and ones that are not differ in what they are *as objects*. That is, objects that are scientific representations possess a certain property, which objects that are not representations lack.[54] In most general terms, a naturalist account seeks to explain scientific representation as a part of the natural order and to give an account of what it means for something to be a scientific representation in physical terms. As a consequence, the difference between representations and non-representations is understood as one between different kinds of things in the world.

There are two ways to understand this claim. On a narrow understanding, it is an intrinsic property of the object that marks the relevant difference between representations and non-representations. On a broader understanding, this difference can also be rooted in a relational property.

The non-naturalist denies this and holds that there is no physical difference between objects that are scientific representations and ones that are not. On this view, we are looking in vein for a property (intrinsic or relational) that representations possess and non-representations lack simply because there is no such property. What marks the difference between the two classes has nothing to do with what the objects by themselves are. It is factors extrinsic to an object itself (such as the use a scientific community makes of an object, the context in which it is placed, etc.) that turn an object into a scientific representation.

In this section I argue that naturalism is wrong (in both the intrinsic and the relational variety) and that every account that seeks to explain scientific representation in physical terms is doomed to failure right from the start. It is the use we make of an object that turns it into a representation of something else. An object comes to stand for another one only if we take it to do so; nothing is a scientific representation if we do not construe it as such. Scientific representation essentially involves conscious agents that turn something into a representation by what they do with it. For this reason, scientific representation is essentially intentional and talk

[54] This proposal is not committed to the claim that it is the *same* property in every case that marks the difference between representations and non-representations. It is compatible with the belief that different types of representations have to be characterised by different properties. The crucial point on this view is that there exists such a property – whatever it may be – in every single case and that therefore the difference between objects that are representations and ones that are not is reducible to the possession of this property.

about a 'natural' relation holding between a model and its target is out of place. There simply is no such thing as 'scientific representation in nature'.[55]

Before providing arguments for this claim, let me state two consequences and two provisos. The first consequence of the denial that scientific representation can be naturalised is that there are no in-principle restrictions on the kinds of objects that could be used as scientific representations; that is, there is no fundamental distinction between objects that could function representationally and ones that cannot. This is not to say that every object is equally useful or equally suited to serve as a scientific representation; nor is it to say that because everything could be a representation that everything actually is. The point simply is that there is no in-principle reason to preclude an object, however outlandish it may seem, from being used as a scientific representation.

The second consequence is of heuristic character. The denial that there can be a naturalistic account of scientific representation shifts the focus from the question of what scientific representations *are* to the consideration of the human activity of *scientifically representing.* What a theory of scientific representation primarily has to account for is not what kinds of objects representations are. Rather it has to come to terms with the question of what kinds of actions on the side of the user turn an object into a scientific representation. What does a scientist (or a scientific community) have to do in order to use an object in a representational way? Or to put it another way, what conditions have to fall in place for something to be a scientific representation of something else? This is not a trivial question because the denial that representations are part of the natural order of things does not imply that an act of sheer stipulation turns anything you like into a representation. It would be a faulty conclusion that a non-naturalist stance implies that there is nothing further to say about representation than that it is constituted by an act of declaration. There are no representations by *fiat*! For sure, such an act of declaration is a starting point, but no more than that. This leaves us with the question of what further conditions have to fall into place for successful scientific representation. I address this question in the next chapter.

---

[55] A view similar to this is stated in Wartofsky (1979, xviii-xxii), but without arguments to support it.

The first proviso I would like to add is that I strictly limit the claim that representation cannot be naturalised to *scientific* representation. What I say in this section does not rule out that other forms of representation – in particular mental representation – could, at least in principle, be naturalised; I don't take a stance on this issue. Moreover, a naturalised view of mental representation, for instance, is perfectly compatible with the non-naturalisation claim that I put forward with respect to scientific representation. What I claim is that there is no natural relation between a scientific model and its target and that the required representational relation essentially depends on the actions of a conscious being. This, however, does not presuppose any view on what consciousness (or intentionality) is.

Second, naturalism is a stance rather than a narrowly defined doctrine and as such it has many different faces in different contexts. For this reason I should emphasise that I only reject one particular brand of naturalism – the one defined above – and not naturalism *per se*. In fact, I will put forward a certain kind of methodological naturalism in the next chapter.

This said, why should we believe that scientific representation cannot be naturalised? I first present some 'heuristic evidence' why we seem to be on the safe side when rejecting naturalism and then offer a general argument for the conclusion that naturalisation necessarily fails.

To begin with, it is worth observing that all currently known attempts at naturalisation fail when it comes to scientific representation. This is either because they can be shown not to succeed or because the basic ideas underlying the naturalisation have been devised in a different context (mental or pictorial representation) and cannot be carried over to scientific representation.

As to the first group, it is worth noticing that isomorphism and similarity accounts *could* be viewed as attempts at naturalising scientific representation – though I should be careful to point out that their proponents have never explicitly endorsed such a programme.[56] To say that scientific representation can be analysed in terms of either similarity or isomorphism can be understood as the claim that scientific representation is reducible to an objective relationship that holds between model and target. However, this does not work, as I have argued Part II. Neither

[56] Giere (2002) and van Fraassen (1997) are explicit about the fact that they do not endorse such a programme.

isomorphism nor similarity is sufficient to establish the appropriate representational relation between model and target.

As far as naturalisation techniques that are not applicable to scientific representation are concerned, four suggestions come to mind: causal connection, covariance, teleology, and functional roles. Let me consider these briefly one at a time.

With regards to photographs it has been suggested that they represent what they do by being brought about in a certain way; that is, by having the right causal connection to the thing depicted. On this suggestion, a photograph is a depiction of the 'original scene' – the thing the camera was pointed at – because the colour pattern we see on the little shiny piece of paper in front of us results from a camera being pointed at a scene, thus allowing certain rays of light to fall upon a photo-sensitive film, which after having undergone some chemical processes transforms into the picture we see. In short, it is the aetiology of the representing device that furnishes the explanation for its representational character.[57]

This suggestion, whatever its value in the case of photographs, is of no help to naturalising scientific representation simply because models are not photographs, nor are they produced in a way that is similar to the production of a photograph. We simply don't point cameras – or any other recording device – at atoms, populations or markets to obtain a model of these. A model is a construct invented by a scientist to achieve certain goals and as such it is a creation of her spirit, provoked but not caused by nature.

Within the philosophy of mind, one of the major attempts at naturalising representation are so-called covariance theories. On these views, the fact that R represents T is grounded in the fact that the occurrence of R covaries with that of T. For instance, the firing of a neural structure in the visual system is said to represent a dog if its firing covaries with the occurrence of a dog in the visual field.

This is not true in the case of scientific representation: the occurrence of a certain phenomenon does not covary with the presence of a certain model. On the one hand, there are often many different models of the same phenomenon – just think about the wealth of models of unemployment – which need not all come to mind if

[57] This view is discussed, but not endorsed, in Black (1970, 100-104) and Pitkänen (1981, Ch. 7)

we face the phenomenon. Unlike the firing of a neural structure in the visual system, a scientific model does not necessarily occur to us if the phenomenon appears. On the other hand, a scientific representation can be present even if the phenomenon is absent. Even if an economy develops in a way that there is no more unemployment, there can still be research on the topic using corresponding models. In short, scientific models and their targets do not covary.

Teleological theories expand on this idea by adding to the clause that R represents T by dint of covariation that the two do not only covary as a matter of fact, but must do so because they are biologically *supposed* to. That is, a mental state has a certain content p – it represents p – only if the belief forming mechanism, which produces this state has the function (purpose) to produce it only when p is the case. Different teleological theories then differ depending on the theory of biological function or purpose they adopt.

It is obvious that teleological accounts do not fare better than covariance theories when it comes to scientific representation. First, the above arguments against covariance can equally be directed against teleological theories. Second, whatever appeal the use of biological functions may have in the case of mental content, it seems rather out of place in the case of scientific representation. No scientific model comes to mind as a matter of biological necessity when a phenomenon is present.

The last item in line are functional role theories. These theories posit that R represents T in virtue of the functional role R has in the representational system of a certain agent; that is, R's representational power depends on the relations imposed by certain cognitive processes involving R and other representations in the system. Theories of that kind depart from the observation that concepts seem definable only in conjunction with one another. For example, when we learn the concepts of mechanics – trajectory, force, mass, energy, momentum, etc. – we do not learn independent definitions for each of these, rather we learn how they relate. There are no definitions outside this circle. The functional role theory of representation takes this observation to its extreme and posits that being part of a functional web of this sort is what makes an item representational.

Again, this idea does not carry over to scientific representation. Unlike our beliefs about trajectories and forces, models do not belong to a holistic web of belief. They are not part of an inferential pattern as, for instance, the concepts 'being a force' and 'having momentum' are. Models just are not the kind of things that can figure as the antecedent or the consequent in conditional statements, say. For this reason, trying to fit models into the framework of functional role theories would be something like a category mistake. Of course, different models can operate against a background of shared beliefs, but this is an altogether different matter.

In sum, none of the currently available strategies for naturalising representation is successful in the case of scientific representation. And this is more than just bad luck. On the contrary, every attempt at naturalising scientific representation must fail. In the remainder of this chapter I argue why I think that this is so. To this end I use the method of indiscernible counterparts, which has originally been devised by Arthur Danto (1981) to show that being a piece of art is a status that is conferred on an object independently from what the object as a physical entity is.[58]

To begin with, consider one of Danto's telling examples (1981, 1-3). He invites us to visit a little exhibition consisting of the following pieces. A square of red paint, which is intended by the artist to show the Israelites crossing the Red Sea. Next to it is another painting, exactly like it, by a Danish portraitist called 'Kierkegaard's Mood'. Then there are two other red squares, again exactly like the Israelites crossing the Red Sea. Both are entitled 'Red Square'; one is a clever bit of Moscow landscape, the other a minimalist example of geometrical art. Next in line is another red square of the same sort now called 'Nirvana', a metaphysical painting based on the artists knowledge that the Nirvanic and Samsara orders are identical and the Samsara word is fondly called the Red Dust by its deprecators. To its left there is a still life by an embittered disciple of Matisse, entitiled 'Red Table Cloth', which looks, again, exactly like the other pictures of the exhibition. The last two pieces in this little collection are a canvass, grounded in red lead, upon which, had he lived to execute it, Giorgone would have painted his unrealised masterpiece 'Conversazione Sacra' and a plain red square, just a thing with paint upon it, which is a mere artefact and not a work of art at all.

This exhibition would be rather monotonous since all pieces look exactly the same. Nevertheless, they are very different works of art, belonging to genres as

[58] I should mention, however, that an idea very similar to the one at work in the method of indiscernible counterparts also underlies Putnam's thought experiment with the ant crawling on the beach that I used at the beginning of chapter two.

different as historical painting, psychological portraiture, landscape painting, geometrical abstraction, religious art, and still life; and the exhibition even contains two canvasses that are not pieces of art at all. Needless to say, I cannot revisit Danto's sophisticated discussion of art here, but his main point transpires quite straightforwardly: a work of art *as a work of art* has great many properties of an altogether different sort than those belonging to physical objects materially indistinguishable from them. In the above example we cannot tell the landscape painting from the psychological portrait merely by looking at it. And worse still, we cannot even tell the difference between something that is a piece of art and something that is just a thing with paint upon it by merely considering the physical properties of the objects. The two are, by assumption, materially indiscernible and yet one is a piece of art and the other is not. For this reason, the difference between an object that is a piece of art and one that is not has to be grounded in something other than in what they are *as objects*, i.e. in something other than their physical properties. Danto identifies interpretation as the 'missing element': 'An object *o* is then an artwork only under an interpretation *I*, where *I* is a sort of function that transfigures *o* into *a work*: *I*(*o*)=W. Then even if *o* is a perceptual constant, variations in *I* constitute different works.' (1981, 125).[59]

What constitutes an interpretation and under what conditions it turns an object successfully into an artwork is an entangled issue that need not occupy us here. The point that I am getting at is that we equally find indistinguishable counterparts in science and that therefore the problem of drawing a line between objects that are scientific representations and ones that are not exactly parallels the problem of distinguishing between pieces of art and 'mere' things. Scientific representations can have materially indiscernible counterparts that are not representations of any sort in the same way as a piece of art can have a counterpart of that sort. Moreover, two materially identical things can be representations of very different targets. As an example consider the famous hydraulic machine Bill Phillips devised as a representation of a Keynesian economy.[60] If, for some impenetrable reason, a manufacturer of pipes had built exactly the same hydraulic system for the purpose of

---

[59] For those who find the example with the red squares too contrived, Andy Warhol's brillo boxes, Marcel Duchamp's urinal or Casimir Malevich's black square may serve as 'real cases' of this sort.

[60] See Morgan and Boumans (1998) for a historical discussion of this machine.

promoting his company at an industrial fair, it would not be a representation of an economy. In fact it would not be a representation at all; it would merely be a sample of pipes illustrating their quality and level of technical sophistication. Furthermore, it is by no means necessary to use this hydraulic system as a model of an economy. It is possible that at some point someone finds it convenient to use the same set-up of pipes to represent the dynamics of a population or the water supply system of some futuristic city. For this reason, there is nothing in the pipes as such that turns them into a scientific representation, let alone into a representation of something in particular.

This is by no means an exceptional case. Another example of the same kind is Bohr's model of the atom, which basically is the reinterpretation of the model of the solar system as a model of the atom. And stories similar to this can be told of any model. Just think of Maxwell's billiard balls, the beads connected by springs, the bag model of quark confinement, etc. All of these can be taken to be nothing but what they are – billiard balls or beads – and nothing changes in what they inherently are when scientists use them as models of something else.

Some may now want to object that all this may well be true of hum drum examples like billiard balls, but once we move on to mathematical science the problem vanishes. This is wrong. On the contrary, the problem becomes worse. The same pieces of mathematics can be used in many different contexts and it is nothing in the mathematics itself that determines this use. The examples for the multiple realisability of structure I have given in Chapter 2 are cases in point.

The upshot of this is that the existence of indiscernible counterparts shows that scientific representation cannot be naturalised. The naturalisation thesis has it that scientific representation is (or is reducible to) a physical property of the object (which, to repeat, can be intrinsic or relational). This trivially implies that something that is a scientific representation possesses at least one physical property that a non-representational counterpart does not possess. This, however, stands in contradiction to the above observation that there are materially indiscernible objects one of which is a scientific representation and the other is not. So we cannot believe both that there are such counterparts and that naturalisation is possible. Since I take the former to be a matter of fact, we have to give up the latter. Hence, scientific representation cannot be naturalised in the sense specified above.

It is worth pointing out that this line of argument equally refutes both brands of naturalism. It is obvious that the argument shows that the narrow version – the one that seeks to explain the difference between representations and non-representation in terms of intrinsic properties – is untenable. But the broad version does not fare better. The red squares in our little exhibition are not only intrinsically equivalent, they also bear the same physical relationships to everything else in the world.[61] But if all the squares enter into the same web of physical relations, then it cannot be these relations that mark the difference between the representation of, say, Kierkegaard's mood and the representation of the Moscow landscape. And this argument carries over to scientific models one-to-one. If Phillips' pipe system bears the same physical relations to all other objects in the world when it is a representation of a Keynesian economy as it does when it is a representation of a water supply system or no representation at all, then it cannot be these relations that turn the object into a representation, let alone a representation of one thing rather than another.

For this reason, something other than what objects inherently are must be constitutive of scientific representation; and this is equally true for material and imagined models because the above argument does not make any assumption about the ontology of the model. This, I think, leaves us only with one option: it the use we make of an object that turns it into a representation (where I understand 'use' in the widest sense possible). Hence, scientific representation is an essentially intentional concept. This still leaves us with the question of where representation comes from. Just intending to use something as a representation is not sufficient to turn something

[61] A restriction is needed here. This is not true of spatial properties because no two pictures can occupy the same space. However, this seems besides the point because it is certainly not the difference in spatial arrangement in the exhibition that makes one red square represent Kierkegaard's mood and the other one the Moscow landscape. Further, another possible objection needs to be mentioned. One could argue that different squares have different relations to the rest of the world because they have different histories of production. This, however, does not threaten the above line of argument because the thought experiment with the red squares can easily be improved in a way that blocks this objection. Assume all the pieces in the exhibition that are art pieces (that is, we exclude Giorgione's canvass and the canvass that merely is a square) are produced by people belonging to an artist's collective. Since they are conceptual artists they do not assign any importance to producing their pieces themselves and let their workshop do it. For this reason, the pieces have the same history of production. This, however, does not affect the conclusion.

into a representation. As not everything that is touched by an artist turns into a piece of art, not everything that is touched by a scientist becomes a scientific representation. Neither artists nor scientists have magic powers. But what are the conditions that have to fall in place for successful scientific representation? It is this question that I address in the next chapter.

Chapter 7

Facing the Enigma of Representation

1. Introduction

In the last two chapters I argued that models are objects, that their properties matter to the performance of their representational function, that they function cognitively, and that scientific representation cannot be naturalised. This sets the agenda. What we are in need of is an account of scientific representation that provides us with a response to the enigma of representation and does justice to these general insights. It is such an account that I will develop in this chapter. Moreover, I claimed that the difference between imagined and material models is not relevant from a semantic perspective. That is, material and imagined models represent their targets in the same way. Accordingly, the account developed in this chapter has to be such that it equally applies to both types of models.

2. Scientific Representation – The Elements

Scientific representation, I maintain, involves three relations: T-denotation, display and designation. The presence of each one of these is a necessary condition for successful scientific representation. They are not jointly sufficient, however. The conjunction of these conditions is both necessary and sufficient for something to be a representation that functions cognitively. But it need not be the case that all representations of this kind are scientific; that is, it may well be that the class of scientific representations is a proper subclass of the class of cognitively functioning representations. What, if anything at all, sets off scientific representations from other kinds of cognitively relevant representations is an involved issue, one that I cannot deal with here. (In fact, this is no less than the 'semantic version' of Popper's demarcation problem.)

These three conditions provide us with a partial answer to the enigma of representation, which was, to repeat, to determine what fills the blank in '*M* is a scientific representation of *T* iff ___'. On the view I suggest in this chapter, if *M* is a scientific representation of *T* then (1) *M* T-denotes *T*, (2) *M* displays an aspect *A* of *M* along with a feature *F* of this aspect and (3) *A* designates an aspect *B* of *T*. This answer is partial because it only specifies necessary conditions and leaves it open in what way – if at all – they would have to be amended in order to render them sufficient as well.

Heuristic considerations and the overall picture

As in the previous chapters, the analogy with other forms of representation is helpful in understanding models. In the present case, language gives the lead. The first thing we have to realise is that models make claims, in that they are like sentences. The model of the solar system claims that planets move on elliptical orbits or the chain model of a polymer claims that the length of a polymer is a function of its temperature.[62] So when we set out in this chapter to construct a semantics for models and to this end steal a glance at language, the point of reference is sentences rather than individual words.

To make a start, consider a simple situation of the kind first year physics students deal with. We have an inclined plane in the laboratory whose inclination is α and whose surface is even and well polished. Now we put a metal cylinder of mass m_c on this plane whose surface is equally even and polished. What is the magnitude of the force accelerating the cylinder when it moves along the plane? To answer this question we swiftly come up with a simple mechanical model, which consists of the following ingredients. First, we have an inclined plane with inclination α and a cylinder with homogenous mass distribution on it. Second, we take the plane to be frictionless. Third, we assume that the only force in this system is the linearised law of gravitation $G=m*g$, where $g=9.81\text{m/s}^2$ is a constant, m is the mass of any object in

[62] Those who think that only conscious users of a language can make claims can rephrase this point as follows: the model legitimates its user to claim that planets move on elliptical orbits etc.

the force field and '*' stands for the multiplication operation. In this model, the force acting on the block along the plane is $G_p = m_c * g * \sin(\alpha)$.

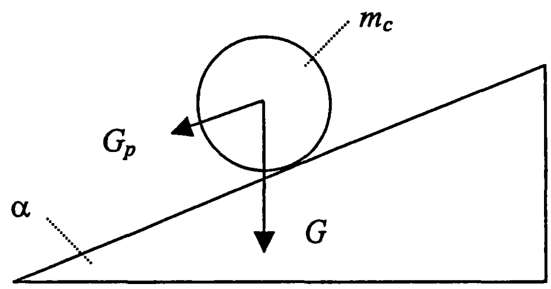

Fig. 5: Schematic picture of the model of the inclined plane.

As it stands, this is a result about the model. But since the model is proffered as a representation of the situation in the laboratory it makes the claim that the force acting on the cylinder when it moves along the plane is more or less G_p in the laboratory as well. How can the model do this?

To answer this question it is helpful to first have a look at its corresponding verbal claim, the sentence 'the force acting on the cylinder along the plane is more or less G_p'. This sentence can be analysed as having the structure '*a* is *P*', where '*a*' is singular term referring to the system in the laboratory and '*P*' is a general term standing for 'having more or less force G_p acting on the cylinder along the plane'. Trivially, '*a* is *P*' asserts that *a* is *P*. Why is this so? Simply because '*a*' refers to the target system, '*P*' stands for a feature of the target, and the auxiliary verb 'is' indicates that *a* possesses *P*.

This said, let us return to the original question, which is, to repeat: how does the model *M* make the claim that the force acting on the cylinder along the plane is more or less G_p? Or to put it another way, how does the model *M* represent the matter of fact that the force acting on the cylinder along the plane is more or less G_p? Given what I have just said about the sentence '*a* is *P*', I suggest we draw the following analogy. The model *as a whole* corresponds to the term *a*, the property *having force G_p acting on the cylinder along the plane* as displayed in the model *M* corresponds to the predicate *P*, and the fact that this property is displayed in the model does the work of the auxiliary verb 'is'. Now we can offer the following tentative account of how the model represents the aforementioned matter of fact. The model *M* as a whole denotes the target system T as a whole (I call this 'T-denotation'), the property *having force* G_p *acting on the cylinder along the plane* as displayed in the model M designates the property *having more or less force* G_p *acting on the cylinder along the plane* of the laboratory system and the fact that the model M displays this property indicates that the target T actually possesses the property.

Two elements of this account need further qualification. First, what exactly do we mean when we say that *having force* G_p *acting on the cylinder along the plane* as displayed in the model M designates the property *having more or less force* G_p *acting on the cylinder along the plane* in the target? As I see it, this claim has to be analysed in several steps as follows. (1) To begin with, we have to identify the aspects that we are dealing with in both the model and the target, and then stipulate that the relevant aspect in the model stands for the relevant aspect in the target. In the present case this is simple. In both the model and the target it is the forces in the system that we are interested in. Call the force aspect of the model A and the force aspect of the target B. Then we stipulate that A stands for B or, more colloquially, that the forces in the model stand for the forces in the target. To say that A stands for B is a synonym for saying that A denotes B. To keep this denotation relation apart from the one that holds between the model as a whole and the target as a whole, I call it 'A-denotation'. This is unproblematic, but it needs to be stated explicitly. The troublesome part of the claim is the qualification 'more or less'. How are we to understand this? What this amounts to, I think, is the following. (2) By assumption, there is a force of magnitude G_p acting on the cylinder along the plane in the model. This is a feature of the force aspect A of the model, call this feature F. (3) Furthermore, there is a force G_p' in the target acting on the cylinder along the plane. This is a feature of the force aspect B of the target; call this feature G. (4) The features F and G relate to each other in the following way: the forces act in the same direction and their magnitues are close to each other by certain accepted standards of closeness. I call this specification of how the features F and G relate the 'link' between the two. This may seem a bit a roundabout way of putting a simple matter, but as we proceed it will become clear why these complications are necessary.

Second, what does it mean for a model M to *display* a certain aspect A with a certain feature F? I take this to involve two things. (1) The model has to posses (or instantiate) A and F. (2) Loosely speaking, A and F have to be considered relevant to a certain problem by the scientist using the model. A model may possess many aspects and features we are not interested in. So we have to select which ones are relevant. I call this selection process 'thematisation'. An aspect and a feature that are both possessed by the model and thematised are displayed.

Now we have gathered together the basic elements of the account of scientific representation I suggest. In general, a model *M* makes the claim that the target *T* has feature *G* (or: it represents the fact that *T* has *G*) iff (1) *M* T-denotes *T*, (2) *M* displays aspect *A* of which *F* is a feature, meaning that both and *A* and *F* are possessed and thematised, and (3) *A* designates *B*, meaning that *A* A-denotes *B* and that *F* is linked to *G*. If we finally also take into account that an aspect characteristically has several features, we obtain the following picture:

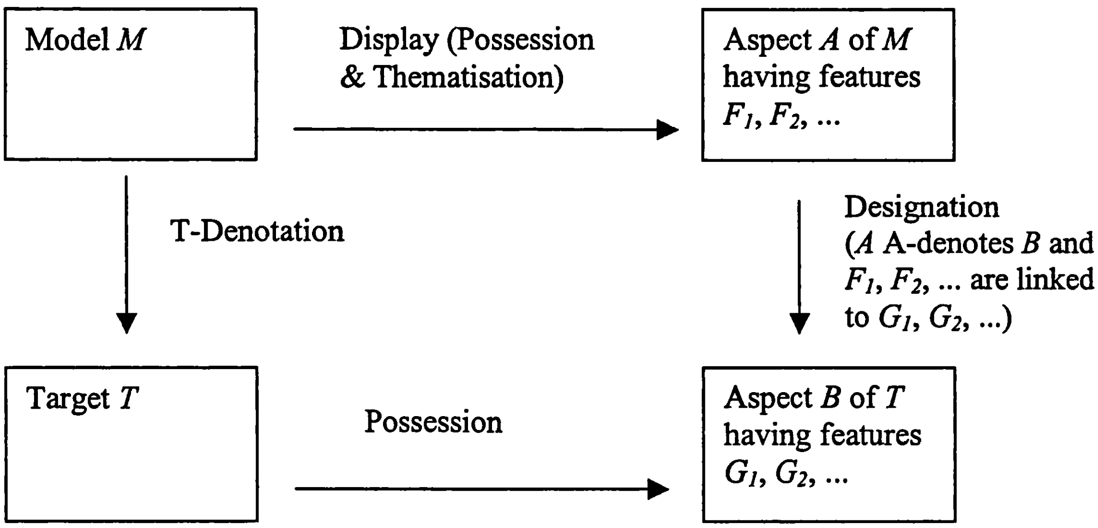

Fig. 6: A three-fold conception of scientific representation

Now that we have the grand picture a hand, let me discuss the above example again. This is to further illustrate how the account works and to remove the impression that it contains a lot of unnecessary detail. There are three aspects in the model. The first two are simply its parts: the plane and the cylinder. The third aspect is the forces acting in the system. Below is a list of these aspects along with the features they have.

Model	**Target**
1^{st} Aspect: the cylinder in *M*	1^{st} Aspect: the cylinder in *T*
Features:	Features:
Spherical cross-section	Nearly spherical cross-section
Homogenous mass distribution	Nearly homogenous mass distribution
Total mass *m*	Total mass *m*
2^{nd} Aspect: the plane in *M*	2^{nd} Aspect: the plane in *T*
Features:	Features:
Surface is a perfect plane	Surface is a nearly perfect plane
Inclination α	Inclination α
frictionless	not much but some friction
3^{rd} Aspect: the forces in *M*	3^{rd} Aspect: the forces in *T*
Features:	Features:
Linearised Gravitation	Newtonian gravitation ($1/r^2$)
Force along the plane:	Force along the plane:
$G_p = m_c * g * \sin(\alpha)$	$G_p' = m_c * g * \sin(\alpha)$ + friction + air resistance + ...

In order for the model to represent the target, the following conditions have to fall in place. (1) The model has to T-refer to the target; that is, the model has to stand for its target system. (2) The aspects and features salient to the problem (the ones listed on the left hand side) at hand need to be possessed by the model and they need to be thematised. (3) Aspects in the model have to stand for (that is, A-denote) aspects of the target. In the present case this is simple (it need not always be that simple): the cylinder in *M* stands for the cylinder in *T*, and so do the planes and the forces. Then we have to link the features, which we have to do one by one. We have to specify how the cross-section of the cylinder in the model relates to the one of the cylinder in the target, how the mass distribution of the cylinder in the model relates to the one of the cylinder in the target, and so on. If all that is done, then M is a scientific representation of T.

This sets the agenda. First, I should present a more detailed account of the three basic relations – T-reference, display and designation – and of how they work together; and I have to say more about what I mean by aspects and features. Then I have to argue that this picture, which I extracted from a simple example, covers more complex cases of scientific modelling as well. Finally, I have to explain how this picture accounts for the fact that we can learn from models. In Chapter 6 I observed that this is one of the defining features of scientific representations but from what I have said so far it has not yet become clear how the suggested view can account for this fact. This is the project for this chapter.

T-Denotation

The first constitutive aspect of representation is T-denotation (or reference, I use these terms interchangeably). A model stands for its target and it does so by denoting it; to recognise something as a model implies recognising that it functions referentially.[63] A model of the solar system refers to the solar system, a model of the hydrogen atom refers to hydrogen atoms, and a mechanical model of a toboggan sliding down a hill refers to toboggans sliding down hills. The prefix 'T' (for 'target') is added to 'denotation' to forestall confusion. Scientific representation involves two different kinds of denotation. On the one hand, there is the T-denotation relation just mentioned in which the model as a whole and the target as a whole enter. On the other hand, the current analysis of scientific representation involves a second kind of denotation, which holds between an aspect of the model and an aspect of the target. These are different relations that can obtain independently of one another and we should be careful to keep them apart.

What is the nature of the referential relation between model and target, and under what circumstances is it true that the model refers to the target? It is a consequence of the fact that scientific representation cannot be naturalised that reference is not fixed by what either the model or the target inherently are, nor is it fixed by any physical relation in which the two enter. To search for the specific

[63] In this I follow Goodman (1968, Ch. 1), who placed denotation at the heart of representation.

properties a model has to instantiate in order to refer to its target is a non-starter simply because there aren't any. Anything can, at least in principle, refer to anything else. This is to say that the referential relation between model and target is no different from the one between a name and its bearer (or a predicate and the objects in its extension, I come to this later on). That a couple of beads connected by springs stand for a polymer no more depends on any property these beads possess than the fact that the expression 'tree' refers to trees depends on any matter of fact about the word 'tree', such as being composed of four letters or being formed in accordance with English orthography. There are almost as many words that refer to trees as there are languages and it is a matter of convention which one we use. Reference is a relation between a symbol and the objects to which it applies, and nothing in that hinges on whether the symbol is a verbal expression of any sort or a scientific model. Or to put it another way, reference is an external relation regardless of whether the representative devices are words or models.

For this reason, the question of under what condition a model refers to its target has to be discussed along the lines of the question of under what conditions linguistic expressions refer to their referents. At bottom, the question of why the term 'water' denotes H_2O exactly parallels the question of why a model of the water molecule T-denotes H_2O.

That is about as much as one can (and should!) say about this issue within a general account of representation. This is for the following reasons. No doubt, reference is a real phenomenon; we use words and we succeed in referring to things in the world. Yet how this is possible and how we succeed in doing so is less obvious. In virtue of what does a term denote? Answering this question requires a theory that explains the relation a term bears to its referent. This is where all the screaming and shouting begins. The problem of reference is an extremely convoluted topic and to date no generally accepted theory has emerged from the debates. For this reason it does not seem advisable to get into the muddle of theories of reference in the context of an outline of theory of scientific representation. But nothing is lost for such a theory by evading the topic, because nothing in the theory hinges on how reference is established. The only thing that matters is *that* reference is established, and this, as I mentioned, is hardly a bone of contention. Needless to say, there may be particular cases in which reference is dubious, or at least controversial (superstrings, fibre bundles, the gene for alcoholism, fitness of an organism, the invisible hand, etc.), and there are other cases in which, given our present knowledge, it fails (phlogiston, ether, etc.). But this does not discredit reference as such; it merely reminds us to be on our guard. In short, a theory of scientific representation does not need to wait for the development of adequate theories of reference in order to proceed; it is enough to know *that* it can be done, and there is little doubt about this.[64]

A further complication is that even if reference were a neat matter, we would need several theories of reference rather than only one. The things models can refer to are so varied that a single theory may not suffice. Some targets are particulars (the solar system, the universe, the surface of the Sun, etc.), others are types (the hydrogen atom, polymers, ecosystems, markets, etc.). Some are objects of our immediate acquaintance (pendula, bridges, predators and preys), others, though still macroscopic, are too far away or too big for us to have immediate experiences of (other galaxies, black holes, etc.), and yet others are too small for us to see (atoms, quarks, etc.). Each of these cases bears its own difficulties and it is doubtful, to say the least, that there is one theory of reference that will cover all of them.

Last but not least, there is a further aspect supporting the view that problems in connection with reference are not genuine problems of a theory of scientific representation. As the case studies below illustrate, we quite often borrow the reference of models from words. In the model of the Sun-Earth system, which I discuss in detail below, we establish reference simply by using the corresponding verbal expressions.

In sum, how to establish T-reference is a problematic and convoluted issue. But since reference as a phenomenon is very real and disagreement about the subject matter concerns the *how* and not the *that*, I suggest leaving it there for the time being and focusing on the problems that are peculiar to a theory of scientific representation.

Before discussing the other elements I take to be constitutive of scientific representation I would like to address a possible criticism. In Chapter 1 I argued that representation does not presuppose realism. However, so the objection goes, by

[64] This point has also been made by Giere (1985, 77), in a different context however.

requiring that denotation is established my suggestion actually does presuppose realism, which is contradictory.

This objection is best addressed by qualifying the notion of realism that I deny that representation presupposes. To represent an item does not amount to giving a mirror image, or to make a copy of that item. A representation can be alike to its target, but it does not have to be. There is nothing in the notion of a representation that ties it to imitation or copying. It is this sense that I deny that representation presupposes realism.

The notion of representation does, however, presuppose that there exists something that a representation is a representation of. On my view, an account of representation (at least within the context of science) cannot do away with the existence of the target; by talking about '*re*presentation' we presuppose that there is something to be *re*presented. A model can only be a representation of the hydrogen atom, say, if there is something that corresponds to the model. This thing does not have to have all the properties ascribed to it in the model but it cannot be 'nothing'; if there is no target system there is no representation either. Consider for instance one of Maxwell's models of the ether. I take it that we would not want to say that this model represents the ether simply because there is, as far as we know, no ether. In cases like this it seems more appropriate to talk about 'presentation' rather than 'representation'. The ether model presents a fictional entity to us, but one of which we now do not believe that it exists. A representation can be inaccurate, stylised or give us a picture of the target that is distorted in many other ways; but it has to represent something that exists. In this sense representation presupposes realism. Moreover, this does not change when we deal with unobservables. The problem with unobservables is epistemic, not semantic. How we know that unobservable entities exist is a time-honoured problem, but not one that affects semantics. If a model is to represent an unobservable entity we assume that this entity exists. How we come to know that it does is just a different issue.

Display

Before I discuss display, I should clarify my use of the terms 'aspect' and 'feature'. They are vague, and deliberately so, because the account should be as flexible as possible to make room for different sorts of representation. Rather than venturing a general definition of 'aspect' and 'feature' (which I doubt that there is), let me illustrate what I have in mind by dint of some examples, all of which will be discussed in greater detail below. First, in some simple cases the aspect is a property (being polluted, for instance) and the feature is the concrete value (the number of pollutants per cubic meter, for instance) the pollution assumes in a certain model. Second, properties can be more complex. In a biological model, for instance, the aspect we are interested in can be the differentiation of cells, and features of this aspect are the details of how certain cells do differentiate in some model organism. Third, mechanical models often consist of clearly identifiable parts. The above model, for instance, consists of an inclined plane and a cylinder. In this case the parts are aspects and the properties they possess (in the case of the plane: having inclination α and being frictionless) are the features. This list is by no means exhaustive, but I think it conveys the main idea of what I have in mind when I talk about 'aspects' and 'features'.

A further preliminary issue needs to be settled. What is it that is being displayed in a model? Aspects, features, or both? The answer is: both. As the above examples show, it is always both the aspects and some of its features that are displayed. There is a slight asymmetry between the two, however. It is, in principle at least, possible that an aspect is displayed without any of its features being displayed as well – although, such a model would be rather useless. The inverse is not possible. A model cannot display a certain feature without also displaying the corresponding aspect. A model cannot, for instance, display the frictionlessness of the plane without also displaying the plane. Why this is so will become clear from the discussion below. To avoid undue verbosity in what follows, I exploit this asymmetry and always only talk about the display of a feature, whereby the display of the aspect to which it belongs is implied. This is merely a matter of presentation of no philosophical import. It would just lead to unreadable sentences if one would always mention both.

This said, let me state the basic definition: a model M displays a feature F iff F is both possessed (or instantiated, I use these terms synonymously) by M and thematised. Let me introduce these in turn.

Possession: In order for a model to display an aspect and some of its features it is a necessary condition that the model possesses this aspect and some of its features. The model of the inclined plane possesses the feature of having accelerating force

G_p; the bob of a model of the pendulum possesses mass; planets in a model of the solar system possess shape and a certain geological constitution; the model of a molecule possess a certain geometrical structure along with some dynamical properties; the model of a bridge instantiates a spatial arrangement; the model of an agent instantiates her preference structure; and so on. This is perfectly possible since, as I have argued in Chapter 5, models are objects, either imagined or physical.

Thematisation:[65] Possession is necessary but not sufficient for display. A model can possess a feature without displaying it. Maxwell's billiard balls display their dynamical properties and but not their internal mass distribution. Display is selective in that not every possessed feature needs to be displayed. This raises the question of what marks the difference between possessed features that are displayed and ones that are not. There is some temptation to answer that the features displayed are the ones that the item shares with its target. This is mistaken. First, the properties the model displays need not be identical to the ones the target possesses (recall the frictionless plane). Second, there may be properties the model and the target share but which the model does not display. Beads on springs can have the same average mass density, say, as the polymer but this property is not displayed. Clearly, joint possession does not do.

Display, unlike possession, is not intrinsic. There is nothing in the model as an object that distinguishes between features that are displayed and ones that are merely possessed. The relevant difference lies in the use we make of the model. When we take a certain entity to be a model of something else we select some among the many features the entity possesses to become the focus of our attention because we think that they are relevant to the investigation we want to carry out. To learn about the target, scientists have to concentrate on a certain narrow array of features in the model that seem useful in dealing with the particular problem the model is intended to tackle. Borrowing a term of Wollheim's (1987, 20), I refer to this selection of relevant properties as *thematisation*. It is a necessary condition for a feature to be displayed that it is thematised. For this reason, there is a pragmatic aspect to display. What features an item displays depends on its function; and the same item can

[65] My discussion of thematisation owes a lot to Goodman and Elgin's discussion of exemplification; see Goodman (1968, Ch. 2), Elgin (1983, Ch. 5; 1996, Ch. 6).

perform a variety of functions in different contexts. It is up to the user to decide which of the properties an item possesses are displayed.

Let me now briefly dispel some possible worries that may have come up in connection with the display of features, and then draw some consequences. So far I have been playing fast and loose with the term 'feature'. What do I mean by it? Whatever you want, really. A model can display almost anything: properties, relations, structures, attributes, patterns, qualities, facts, and so forth. I use 'feature' as an umbrella term to cover all these cases. The term is sufficiently unspecific to make room for a variety of options but still effectively conveys the central idea. There really are no restrictions as to what features a model can display; and if there are restrictions of that sort, they are imposed by the needs and methodological bounds of a particular scientific context, and not by any philosophical view on representation. What really matters is that a characteristic of the model, whatever this characteristic may be, is matched with a characteristic of the target and that a link between the two is specified. Whether this correspondence is property to property, part to part, relation to relation, or what have you does not matter from a representation-theoretic point of view. Where 'feature' talk seems to impose constraints on what we can do we have to replace it by something else, for instance talk about 'relations' or 'properties' of a model.

Two related matters deserve brief mention. First, there have been extended debates on the nature of properties dealing with questions such as whether properties are universals or tropes, whether they persist through time, whether they must be determinates, or whether they have to be instantiated. Though interesting in their own right, in the context at hand these issues have no importance. Furthermore, in a nominalistic account of display, the talk of properties is replaced by talk of labels and predicates. The nominalist can accept property talk as a harmless matter of convenience that has to be substituted by talk of predicates or labels if greater rigour is necessary (see for instance Goodman 1968, 54-67).

Second, it is a much debated issue whether a predicate refers to a property, a class, or each and every individual in the class. Does 'beautiful' refer to the property beauty, to the class of beautiful things or each beautiful item individually? Again, my answer is deflationary. For the needs of a theory of representation it does not matter. I mainly stick to the first option in what follows; but this choice is dictated by convenience rather than by the adoption of a particular metaphysical stance on properties. The problem of linking, which is the main subject of the next section, is more easily stated and discussed when using property talk. But nothing depends on it. Everything I say about linking can be recast in the idiom of classes or individuals without any loss.

In sum, in the context of a theory of representation, there is no need to worry about the nature of the features an item displays.

Let me now draw attention to three important features of display. First, the features a model displays are epistemically accessible; the model presents them in a way contrived to render them salient. This may involve filtering out impurities, removing unwanted irregularities, or presenting them in an unusual setting. On the extreme end, some models in both experimental and theoretical contexts are designed such that they screen off all but one factor in order to render its operation palpable, which is not accessible when the factor occurs mixed with other factors (this is of particular importance in causal modelling). Many models in economics (e.g. the Phillips curve), biology (e.g. predator-prey models studying the interaction of two species in a stable environment), as well as in physics (e.g. models of the interaction of particles neglecting all but one force acting between them) are of this sort.

Second, display is an intentional notion. Thematisation is always for an end. An agent thematises a certain feature in pursuit of a purpose. It is the imposition of a purpose upon the model that determines what will be thematised. If we are concerned with the development of cells, we have to thematise the pattern of cell differentiation we find in an organism. That this organism may also possess many other features, for instance that it has a certain metabolism, is irrelevant unless it is doing so in connection with the features that interest us.

Third, display is highly context sensitive. Models operate against a constellation of explicit or tacit assumptions and an agent ignorant of these will be unable to use the model. With the change of these assumptions the model can come to display other features. Furthermore, what features a model displays depends on its function; and the same object can perform a variety of different functions. This is quite apparent when we change the scientific context. Just think of the standard lab mouse. This mouse is used in many different experiments: stomach cancer, skin diseases, immune reactions, nutrition, and so on. In each of these the mouse *displays* different features, although it always *possesses* the same features. In one context it displays certain properties of its skin, in the other certain patterns of vitamin absorption in its digestive system. Or take a diamond. In one context it displays hardness, in another one a certain refraction behaviour of light. In short, the properties an object displays vary with our interest.

Designation

An aspect A displayed in the model designates an aspect B of the target iff A A-denotes B and some of the features F_i of A are linked to some of the features G_i of B. There are three elements in this definition that need explanation: A-denotation, linking, and the fact that the designating aspect has to be displayed.

The designating aspect has to be displayed: The designation relation does not hold between an aspect A *per se* and another aspect B; it holds between A *as displayed in the model* and B. This is because designation is context sensitive. A certain aspect of a model can stand for different target aspects in different contexts. For instance, when used as a mechanical model, the elongation of a model of a pendulum (i.e. the distance of the pendulum bob from its equilibrium position) stands for the elongation of the steel ball oscillating on a spring in the laboratory. However, when the same model is used as an analogue of an electric circuit, the elongation stands for the voltage of the target system. For this reason, it is not the aspect A *per se* that stands for B; it is A as instantiated in model M and as thematised in a certain way that does.[66]

A-denotation: It is a necessary condition in order for A designating B that A stands for (denotes) B. In Chapter 6 I argued that the properties models possess as objects matter to how they represent their targets. The most natural way to 'implement' this characteristic in a theory of scientific representation, I think, is to match certain aspects of the model with certain aspects of the target. We then can understand A as 'representative' of B and exploit A to learn about B. How learning takes place is my concern in the next subsection; what matters for now is that

[66] Although this might strike some as peculiar, context dependencies of that sort are quite common. Red as instantiated in the warning light stands for the overheating of the engine; but red *per se* does not. Or green as instantiated in a traffic light stands for the fact that you can move; but green per se does not.

matching an aspect of the model with one of the target minimally involves denotation. *A* has to stand for *B* if we want to use *A* as a 'substitute' for *B*. As I mentioned above, it is important to keep this denotation relation apart from the one that holds between the target and the model (T-denotation). For this reason I add the prefix 'A' (for 'aspect') and call the kind of denotation that is part of designation 'A-denotation'.

When it comes to the question of under what conditions an aspect of the model A-denotes an aspect of the target, the same remarks as the ones made in the case of T-denotation apply. This question also has to be discussed along the lines of under what conditions linguistic expressions refer to their referents. For this reason I won't say more about the problem here.

Link: The crucial ingredient of designation is linking. Models give us knowledge about the target; more specifically, we learn about features of the target by investigating features of the model. For this to be possible there need to be a connection over and above sheer denotation between these properties. Almost anything can denote almost anything else. Learning cannot be explained on the basis of denotation only. What we need to know is in what way the two features relate. For this reason, in addition to denotation we have to provide a specification of how the feature *F* displayed in the model relates to the feature *G* of the target. For want of a better term I call such a specification a 'link'.[67]

In some simple cases, the link between the features of the model and the target is just identity, in others it may be more complicated. Consider again the inclined plane. The model and the target share the features that the inclination of the plane is α and that the mass of the cube is *m*. In these cases the link is identity. This is not true of the other features of the model. The force acting on the mass in the laboratory is not the linearised version of gravitation, it is gravitation which obeys a $1/r^2$ law. To link these two properties amounts to specifying how they relate, which is easy in the present case. Some straightforward algebra soon reveals that the variation of the force as function of the location of the mass within the laboratory is so small that it is beyond measurement precision. So treating the force as constant is a very good

[67] To my knowledge, this has not been put in this way before, but in essence the problem has been recognised. Both Giere's theoretical hypotheses (1988, Ch. 3; 1999, 177ff.) and Morgan's stories (1999b) can be understood as specifications of links.

approximation in the case at hand. But what about friction? Certainly we can observe the difference between a frictionless plane and one that has friction. In this case, one can treat the frictionless plane as an idealisation of the real plane – I come back to idealisation in some detail later on. By assumption, the surfaces of both the plane and the cylinder are well polished so that there is not much friction between the two; and the air resistance is so small that we can neglect it. We then may argue that for the purposes of the present experiment, not much friction is close enough to no friction for the results derived on the assumption of no friction to still be valid.

To find the appropriate links may not always be that easy. On the contrary, in many cases it may be a formidable task. I will come back to the issue in the next section.

Interlude: representation and intentionality

At this point it is worth pointing out that there is a stark contrast as regards the status of intentionality between the view on representation I suggest in this chapter and the structuralist conception. Structuralists by and large[68] think of representation as independent of observers. *M* represents *T* if the two are isomorphic and whether this is the case does not depend on any particular activity of a user, in fact it does not even depend on the presence of a user at all. For this reason, on the structuralist view representation does not involve intentionality. By contrast, the account I suggest in this chapter construes representation as essentially intentional in the sense that what turns an object (physical or imagined) into a representation is the use a scientist or a scientific community makes of it. More to the point, the current account takes representation to be constituted by actions of the user. If there are no users there are no representations; or to put it another way, in a world without conscious beings there are no representations. This is reflected in the use of intentional concepts such as thematisation in the conditions that I take to be necessary for scientific representation.

[68] An exception is van Fraassen (1997).

Learning from models – a two stage approach

There is an intimate connection between knowing and representing. This interdependence is grounded in the fact that models are the units on which significant parts of scientific investigation are carried out. So how does the acquisition of knowledge about the target from its model take place?

We are now in a position to give an answer to this question. Learning from models involves two steps. First, we have to inquire into the features of the model itself and then transfer the findings to the target system by exploiting its links.

The first step consists in getting to know the model itself. This involves, among other things, identifying the relevant features and finding out about the connections between them. This can be done in various ways, depending on what the model is and on what we want to learn from it. In simple cases like the inclined plane this is straightforward. But in many scientific contexts things are less obvious. As Mary Morgan points out (1999a), getting to know the model essentially involves manipulating and using it. Think for instance of Lorenz's famous model of fluid convection. Once the model was available, it was a quite a task to find out what the features of the model were and how they fitted together. Analytic methods failed and so he resorted to computational techniques, which finally led him to the discovery of the so-called butterfly effect.

Examples of this sort abound and not much ingenuity is needed to extend the above list. In all these cases the salient features do not lie bare in a model. On the contrary, finding out what is going on is an appreciable task. What this task involves may vary with the kind of model at issue. But in any case, it does not come for free.

But knowing the model itself does not yet tell us anything about the target. Therefore the second step consists in 'exporting', or 'projecting back', the knowledge we acquire about the model to its corresponding target system. It is at this point that the three relations constitutive of scientific representation come into play. We have to know what the target denotes and what its displayed aspects designate. On the basis of this we can start 'translating' what we know about the model into knowledge about the target system.

In doing so the links between the features of the model and the ones of the target play a crucial role. In the simplest case, the features displayed in the model are identical to the ones possessed by the target. This renders the conversion of knowledge about the model into knowledge about the target trivial. What we know about the model equally holds true of the target. If, for instance, we find out that a model bridge that displays certain architectural features shows a particular resonance behaviour (and we assume that these features are the only factors relevant to resonance), then we can infer that the real bridge, provided it has the same architectural features, exhibits the same resonance behaviour.

This is easy and straightforward; and for this reason scientists often aim at constructing models that function in this way. But despite their desirability, models of this kind are scarce. For the most part, the features of the model and the ones of the target are not identical. Planets are not ideal spheres with spherical mass distributions, polymers are not beads on springs, markets are not free, and ecosystems are not isolated. At this point we have to start reflecting on how model and target relate; that is, we have to give an account of the links between the two. The answer we give to this question determines how the transfer of knowledge is to take place. If, for instance, we have a model we take to be a realistic depiction (it remains to be specified what that means), this transfer is accomplished in a different manner than when we deal with an analogue, or one that involves idealising assumptions. For this reason, the way in which a model represents its target – the links that it has – directly bears on what we can learn from it about physical reality: different ways of representing, different ways of knowing.

There is a further important aspect to learning from models: not all features of the model have the same status. Some are more basic than others. Consider again the above example. The fricionlessness of the plane, the inclination α, the mass m and the force G acting on m are features the model displays by construction. There is nothing we have to find out about that. Furthermore, these features designate features of the target because we directly let the features of the model refer to features of the target and specify the relevant links (identity in the cases of inclination and mass, approximation in the case of force and idealisation in the case of friction). This is not the case for the force G_p. First, we learn that the model itself possess G_p by knowing that it has α, m and G and by using the laws of vector addition. In this sense, knowledge about G_p is derivative. Second, we 'translate' this piece of knowledge about the model into knowledge about the system by relying on the designation of the 'basic' features this fact rests on. It is because the model possesses α, m and G

and because α and m have an identity link to the corresponding features of the target, and G and the friction of the plane have an approximation link to each other that we know that the target possess a property G_p' which stands in an approximation relation to G_p. In this sense, the 'derivative' feature G_p' 'inherits' its designation (and in particular its link) from the 'basic' features α, m and G.

In the next three subsections I discuss simple examples illustrating these claims. Moreover, all three of them involve material models and therefore illustrate the claim, put forward in Chapter 5, that imagined and material models can be covered by the same theory of representation.

A first simple example: samples

When we want to know what the wallpaper on the big reel is like we have a look at the swatch in the shop's pattern book. The swatch possesses exactly the same design as the wallpaper we later buy and therefore we learn by looking at the swatch what design the wallpaper has. In this sense the swatch is a model of the wallpaper. This is so obvious that we do not normally state it explicitly. The swatch has been 'built' in a way to ensure that we can learn from it about its target, namely by simply cutting off a piece from the wallpaper in the stockroom. Models of this sort are commonly referred to as 'samples'.

Samples are important in science as well. When we want to know the proportion of pollutants in a lake, the fraction of ill animals in a population, or the rate of people in a given country infected with a certain virus we take a sample of water, animals, or people and analyse it. This may involve very different things – I will say more about this below. Let me, at this point, just briefly analyse how a sample works in terms of the above notion of representation. For the sake of simplicity, I do this by dint of the wallpaper swatch; the main idea can easily be carried over to more interesting examples. The model as a whole – the swatch we find in the pattern book – denotes its target, the wallpaper on the reel in the shop's or the factory's stock room. This is understood when presenting the swatch *as a sample*. To understand how the swatch functions implies that we take it to refer to the wallpaper on the reel. It also implies that we recognise the design of the swatch as the relevant aspect and thematise it. When we see the swatch we have to focus our attention on the design, because this is what we are interested in. The swatch possess many other properties, which are merely possessed without being thematised (it has a certain average density or a certain chemical constitution, has been produced in a certain place, etc.). Finally, also by being presented with the swatch *as a sample* we know that the relevant aspect of the swatch, its design, stands for the design of the target and that the particular colour pattern the swatch has – the feature of the design – is linked to the colour pattern of wallpaper by identity. Then the swatch warrants the claim that the wallpaper in the stockroom indeed exhibits the pattern we see in the sample, which is true if it is an accurate sample and false if it is not.

A second simple example: measuring air resistance in the wind tunnel

A physical object encounters air resistance when it moves and the magnitude of this resistance strongly depends on the shape of the object. What is the magnitude of the air resistance of a given object? Unfortunately this problem is usually intractable by theoretical means. So scientists resort to experimentation. They build a scale model of the object, i.e. an object that is smaller than the original but has the same shape and surface, put it into a wind tunnel, and experimentally determine the air resistance of this object.

It is clear by now how this case has to be analysed in terms of T-denotation, display and designation. But, and this is the catch, we have to be careful with the link. It would be false to follow the lead of samples and think that the original and the model have the same air resistance because they have the same shape. As a matter of fact, the value of air resistance changes with the square of the scale of the model. For instance, if the model is half of the size of the original, its air resistance is one quarter of the air resistance of the original.[69] For this reason, the link is different in this case than in the case of a sample. With a sample the link is identity; in the present case it is the quadratic scaling law.

[69] This is true only in special cases. Generally, the relation is more complicated. But in any case, it is not identity.

A third simple example: cell differentiation

In biological and medical research it is often the case that one organism is used as a model for another one. The standardised lab mouse is an example for models of this kind. So it is interesting to see how the suggested account of representation works in cases like these.

As a concrete case, I choose a recent success story in biomedical science. Let me begin with some historical background. In the early sixties, Sidney Brenner realised that it was too difficult to study cell differentiation and the development of organs in mammals. But the traditional objects of study of molecular biologists – one-cell organisms such as bacteria or yeast mushrooms – were too simple to reveal anything useful. So he came up with the idea of using the worm *Caenorhabditis elegans* as a model organism for humans (or mammals more generally). This worm is transparent, so that one can directly observe cell division under a microscope. In both humans and worms, all body cells are descendants from the fertilised egg cell. During the development of the embryo the different types of cells, which perform specific functions in certain organs evolve. This allows the study in the worm of certain processes we find in humans. These ideas were taken up by Robert Horvitz and John Sulston who did further work on the worm and the genetic processes involved in cell differentiation. For this research the three were awarded the Nobel Prize in Medicine in the year 2002.

How does this story fit into the framework of the suggested account of representation? First notice that it is not due to an overwhelming interest in simple organisms such as worms that *Caenorhabditis elegans* is studied. If this were the case, the worm would be a sample of its species and its functioning as a model could be explained along the lines of the swatch example above. Rather, *Caenorhabditis elegans* become the subject of intense study but because it reveals certain things about more complex organisms such as humans. This trivially implies that one takes the worm to stand for humans, which amounts to saying that the worm is taken to T-refer to humans.

The worm has a myriad of aspects; but only a few of these are relevant to the question at stake. So one has to choose which aspects are relevant and concentrate on these. This is thematisation. But as the worm is supposed to function as a model for mammals we assume that these possess a set of related aspects for which the aspects of the worm stand. This is A-denotation. Now we need links between certain features. Needless to say, worms are not humans and therefore the two sets of features will not be the identical; so identity links will not do. Still, there must be some similarities between them, otherwise one would not have studied the worm in the first place. What these are is a question biologists have to answer. Once we have a precise specification of this relationship, one can 'translate' knowledge gained in the worm into knowledge about mammals. But it is obvious that without a specification of how the two sets of properties relate it is not clear what we learn about mammals from the worm. So we cannot do without links.

3. Linking

Linking, methodological naturalism, and the problem of quomodity

Linking amounts to specifying in what relation the features the model displays and the ones the target possesses stand. There are two aspects to this problem. First, we have to specify in what relation a certain feature of the model and its counterpart in the target enter; one may call this the semantic problem. Second, a justification for the specification given is needed; this is the epistemic problem. For instance, we claim that the pattern the swatch displays is identical to the one the wallpaper instantiates (semantic aspect) and we justify this contention by pointing to the history of production of the sample (epistemic aspect). In what follows I am concerned with the former rather than the latter – the topic of this thesis is semantics – but I nevertheless make some remarks about the epistemic aspect of linking.

Identity of the features the model displays and the ones the target possesses is certainly the most convenient option because it affords us immediate access to what we are ultimately interested in, namely the properties of the target. Though desirable, identity is rather rare. In fact, at least as far as the more theoretical parts of science are concerned, identity is virtually unattainable. It is just in the most exceptional circumstances that the model possesses *exactly* the same properties as the system. Planes are not frictionless, planets are not spheres, gases are not ideal, the plates of condensers are not infinitely extended, systems are not closed, real agents are not perfectly rational, markets are not in equilibrium, pendulums are not free of dissipation, and the molecules in the wall of a blackbody are not harmonic oscillators, and so on. Distortions of this sort are the norm in science and examples of models displaying properties that their respective targets do not possess abound. And this is not indicative of poor science. On the contrary, often it is falsities of this kind that make many models good models. How then are these to be understood?

One might be tempted to say that this difficulty can be overcome simply by relaxing the identity requirement and taking similarity to be the appropriate relationship between model and target. This will not do. From what I have said about similarity in Chapter 4 it becomes clear that to say that a model and its target have to be similar is no more than a restatement of the problem. Similarity is an abstract term, a blank to be filled, and to fill it is exactly what the problem of linking amounts to.

How are we to respond to this problem? I think we have to recognise that an answer to this problem has to be taxonomic and that there is no way around identifying different representational strategies we take to be scientifically acceptable and then working through every case in its own terms. At the end, this results in a 'dictionary', or catalogue, of representational strategies listing the different strategies and explaining how each of them works. To come up with such a dictionary is a proper research project, one that I cannot possibly embark on here. However, to remove the air of discomfort caused by leaving everything in the open, I briefly discuss two of the important strategies below, namely identity and idealisation. The purpose of these discussions, I should emphasise, is to convey the idea of what an entry in such a dictionary might look like rather than to make substantial advances in the understanding of these strategies. Much effort has been expended in the past in particular on developing theories of idealisation and it is beyond the scope of this chapter to provide a comprehensive survey, let alone a critical discussion of the various suggestions. Before discussing identity and idealisation, three remarks concerning the character of this catalogue need to be made.

First, we have to realise that such a catalogue can never be complete (or if it is we cannot know that it is). There may be ways to understand the link between two features other than the ones we are currently familiar with. The best we can do is to enumerate and discuss some of the currently known options, thereby always keeping in mind that this list is tentative and not exhaustive. Nor are the different options mutually exclusive. In some models several different strategies can be at work at the same time, or the same model can be understood in more than one way.

Second, and most importantly, we cannot expect a philosophical analysis of the different modes of representation to provide us with criteria of when a certain mode of representation is at work and when it is not. That is, we cannot expect recipes to analyse a given model. Such an analysis requires working through each model in its own right. A philosophical analysis has to give us certain notions in terms of which a discussion can be couched; but no 'mechanical decision procedures' can be forthcoming. Moreover, to discuss the issue of what mode of representation is at work in this or that model, for instance, is a scientific and not a philosophical matter. This is because one would expect that in a properly presented model a specification of how it is supposed to represent should be included. But often it is not and many foundational projects in the philosophy of science are concerned with understanding how certain models (or theories for that matter) face reality. For instance, evolutionary models ascribe fitness to organisms. But do real organisms possess fitness? And if so in what way? It is philosophers of biology rather than biologists who are concerned with this question. Or consider models of self-organising criticality. These models ascribe a certain kind of complex behaviour to the entities in the model. But do real objects exhibit the same kind of complex behaviour. Although scientific in nature, this question does not receive the attention it deserves in the scientific literature and it has become a philosophical project to discuss this issue (see my 2002). But these examples notwithstanding, at bottom it is scientists and not philosophers who have to answer the question of how a given model represents.

This methodological naturalism is closely related to the non-naturalisability of scientific representation. If there is nothing in the model as an entity that makes it a representation in the first instance, there certainly is nothing either that makes it a representation of this or that sort. It is the use an object is put to that turns it into a representation; and it is also this use that determines of what kind the representation is. However, to specify how something is used to represent something else is part and parcel of what it means to construct a model – and this a scientific task. Philosophical analysis can help to get clear on what the claims involved in a certain model really are. But in doing so no genuinely philosophical problem is addressed, rather a scientific problem is clarified by philosophers.

Last but not least, I should make a remark about the problem of *quomodity*. It is worth realising that putting together a catalogue of representational strategies of the kind described above amounts to responding to the problem of *quomodity*. Such a catalogue contains the different ways in which aspects of the model can relate to aspects of the target, along with an account of how these ways differ from one another. But specifying different ways in which a model can face its target amounts to nothing less than specifying different modes of representation. And this is the problem of *quomodity* as formulated in Chapter 1.

Identity and Sampling

Identity of the features the model displays and the ones the target possesses is no doubt the most convenient option. It affords us immediate access to what we are ultimately interested in and saves us painstaking translation work. Everything we find out about the model is supposed to carry over to the target without further ado. I shall call items that are used in this way 'identity-type models'.

The most common (but not the only) variety of identity-type models is samples. In sampling we assume that the properties a sample displays are shared by the stuff sampled. The pattern displayed by the wallpaper sample should be the same as the one the wallpaper shows, the reaction to nicotine we find in one mouse should be the same for all mice of the same kind, and so on. What distinguishes samples from other identity-type models is their epistemic aspect, that is, the way in which the identity claim is justified. What makes a sample a sample is its history of production. Roughly, we call something a sample if it is a representative part of the whole it is supposed to stand for. The swatch is a sample of the wallpaper because the little thing we hold in our hands was cut off from the reel in the stockroom from which we buy the rest. At a tasting session, the wine in the glass is a sample of the one in the barrel because it has been taken from there and we assume that all wine in the barrel tastes the same. The same goes for the bit of cheese we cut off from the whole piece to try it before we buy.

Not all samples are that simple. Suppose we want to know the proportion of pollutants in the water of a lake.[70] We take a sample of water from the lake and analyse its chemical composition to learn something about the lake. But not any bottle of water from the lake will do. The concentration of toxins may be greater close to the ground or it may be particularly low near the confluence of the lake and a river. How a sample has to be taken becomes a non-trivial issue. And even more so in the context of social science. How should we select the five hundred or so people to find out the proportion of the population supporting the Labour Party? Going to one end of town rather than the other can produce totally different results.

This raises the question of the fairness of the sample. One might be tempted to reply that the sample is fair if what we find in the sample holds true of the totality. This is true but useless. The whole point of sampling is that we typically don't know what the properties of the totality (the lake, the population, etc.) are and the sample is supposed to tell us. If we had ways to compare the features of the sample directly with those of the relevant totality, sampling would be otiose. So we need standards of fairness independent of direct comparison between the sample and what is sampled.

There is no straightforward answer to the question of where to get such standards from. The crucial aspect is how the sample is taken. But how should we take samples? I take this to be a question of scientific method. In the case of laboratory chemistry the simple advice to stir well might do, while matters in demography are more convoluted. It is up to the methodology of individual sciences to devise methods to obtain unbiased samples.

In other cases sampling amounts to an exercise in induction. A lake or the population of Great Britain are finite totalities and the question is how to take a sample that reflects the 'reality on the ground'. In many cases things are further complicated by the fact that the ensembles at stake are not (or may not be) finite. In medical research doctors look at some humans to find out about all humans, for instance in testing new drugs. But the set of humans is not closed and we cannot possibly know all of them (unlike the population of Britain at a given time; one might, at least in principle, laboriously ask each and every one whether s/he will vote for the Labour Party in the next elections). One examines a few and then infers

[70] This example is due to Goodman and Elgin (1988, 21).

inductively that the results found hold for all the others as well. This is a common strategy in experimental research and the problems it raises are well known.

So much for sampling. Though samples are important, they are not the only kind of identity-type models. A model bridge can have exactly the same configuration as the original bridge or a model of a superconductor can exhibit the same behaviour at low temperatures as the piece of $HgBa_2Ca_2Cu_3O_8$ in the lab. But they are not samples. We do not justify the identity claim by dint of the history of production of the model; it is irrelevant how we construct the model bridge or how we find our model of the superconductor. Rather we invoke background knowledge, engage in experimentation, or point to how the model can be put to use.

Idealisation: ideal limits

If the feature displayed by the target and the one possessed by the model are not identical, the most natural move is to understand the former as an idealisation of the latter. But what does that mean? The problem is that the term 'idealisation' is rather loose and can signify many different things. At the most basic level, an idealisation is a deliberate simplification of something complicated with the objective of making it more tractable. In what follows I discuss one particular way in which this can be achieved, namely ideal limits.

A generic idealising strategy is to 'push to the extreme' a property that a system possesses. As a paradigmatic example consider the ideal gas. At room temperature, the volume occupied by the gas molecules is small compared to the volume the gas occupies; and since the mass of individual molecules is small, their gravitational interaction is small as well. Therefore it seems that a model that takes molecules to be point particles that exert forces on one another only in collision is not too far off the mark because these properties seem to be ideal limits of the properties the gas actually possesses. And similarly with frictionless planes, mass-less strings, spherical planets, and so on. But what exactly are ideal limits?

Two things are needed to render the idea of ideal limits benign: experimental refinements and monotony (Laymon 1987, 1991). First, there must be the possibility of refining actual systems in a way that they are made to approach the postulated limit. With respect to friction, for instance, one has to find a series of experimental refinements that render a tabletop ever more slippery and hence allow real systems to come ever closer to the ideal frictionless surface. These experimental refinements in combination generate a sequence of systems where each successive system more closely approximates the ideal limit (without ever reaching it, however). Second, this sequence has to behave 'correctly': the closer the properties of a system come to the ideal limit, the closer its behaviour has to come to this limit. If we take the motion of a spinning top on a frictionless surface to be the ideal limit of the motion of the same spinning top on a non-frictionless surface, then we have to require that the less friction there is, the closer the motion of the real top comes to the one of the idealised model. Or to put it in more instrumental terms, the closer the real situation comes to the ideal limit, the more accurate the predictions of the model. This is the requirement of monotony. We have a *limit model* if these two conditions fall in place, i.e. if there is a series of experiments approaching the ideal limit in a monotonic way.

This naturally raises the question of what happens if no such sequence is available. The practical problems in producing tabletops coming ever closer to the frictionless ideal may be enormous, but there does not seem to be any in principle problem in doing so. This is not so in other cases.[71] We cannot possibly produce a sequence of systems in which Planck's constant approaches zero. Though formally similar – in either case there is a parameter tending towards zero (the coefficient of friction and Planck's constant) – there seems to be an in-principle difference between these two cases. A system with Planck's constant tending towards zero cannot be understood as an idealisation of a quantum system in the same way a frictionless table is an idealisation of its counterpart with friction. The enormous difficulties we have in understanding the relation between classical and quantum mechanics show how far away from each other these two situations are. Therefore, cases in which no such sequence is available cannot be understood as limit models and have to be treated in a different way.

Yet in other cases it may not be clear whether there are such limits or not. For instance, mathematical knot theory is a branch of topology and as such it deals with one-dimensional strings. But physical strings – bootlaces for instance – have finite width. Hence the question arises whether, and if yes, in what sense the results of

[71] This corresponds to Rohrlich's distinction between factual and counterfactual limits (1989, 1165).

mathematical knot theory carry over to physical situations. This is a currently hotly debated topic and a great effort is made to get a handle on physical strings. From the current perspective these efforts can be interpreted as the attempt to understand the characteristics of the ideal involved.

Similar provisos apply to situations in which a sequence is available but it fails to be monotonic. An idealised model cannot be understood as a limit model if the limit behaves in a way totally different from how the systems in the sequence behave. If the dynamics of a spinning top on a frictionless plane were totally different from the one on a plane with finite friction, the model would be of no use, or in any case not of the use we are making of it.

4. Illustration: Modelling the Sun-Earth System

In this section I discuss the well-known mechanical model of the Sun-Earth system. The purpose of this discussion is to show in detail how the definitions of the suggested account of representation work out in this case. I continue the discussion of this example in the next chapter and show how the model presented in this section serves as a foothold for the mathematical treatment of the problem. I choose this simple and well-known example as an illustration of my claims because it is interesting to realise that even simple cases work in the way I suggest and that one does not have to move into fancy physics to make the account work.

The first step in the introduction of a model consists in specifying what the model itself is. A standard textbook story for the Sun-Earth system is something along the following lines.[72] Think of the Sun-Earth system as one consisting of two spheres with spherical mass distribution, a big one standing for the Sun and a small one standing for the Earth. Then assume that the only force in the system is gravitation acting between the two centres of the spheres and that the Sun is held fixed. Given this, we can use Newtonian mechanics to calculate the trajectory of the Earth.

[72] Specifications of this kind can be found in most basic physics textbooks (for instance Feynman 1963, Secs. 9.7 and 13.4; Ohanian 1985, Ch. 9; and Young and Freedman 2000, Ch. 12).

In order to understand how exactly this model represents its target system we need a more detailed description of the model. In particular, we have to figure out what the aspects and their features are. I suggest the following reconstruction of the usual physics textbook story:

Model of the Sun-Earth system

1st Aspect: body b_1

Features:

$F_1 = b_1$ has mass m_1.

$F_2 = b_1$ has the shape of a perfect sphere.

$F_3 = b_1$ has a spherical mass distribution.[73]

2nd Aspect: body b_2

Features:

$F_4 = b_2$ has mass m_2.

$F_5 = b_2$ has the shape of a perfect sphere.

$F_6 = b_2$ has a spherical mass distribution.

3rd Aspect: forces in the system

Features:

$F_7 =$ The force acting between b_1 and b_2 is $f=g*m_1*m_2/d^2$, where d is the distance between the centres of b_1 and b_2 and g is a constant.

$F_8 =$ There are no other forces acting on either b_1 or b_2.

4th Aspect: b_1 and b_2 are located in space-time

Features:

$F_9 =$ The space-time is classical.

5th Aspect: motion of the bodies

Features:

$F_{10} = b_1$ is held fixed and only b_2 can move.

[73] That is, the mass distribution may vary with the distance from the centre of the sphere but not with the angles. A homogenous mass distribution is a special case of a spherical mass distribution

Display. The aspects and features in the above list are all possessed by construction. Furthermore they are all seminal to what we ultimately want to know, namely the motion of the bodies, and for this reason we thematise them. It might now seem as if thematisation were a trivial matter because all features the model has are thematised as well. This is not true. The model has features we do not thematise. For instance, there is a central force field inside both spheres whose strength depends linearly on the distance to the centre (this is a consequence of the fact that the mass distributions are spherical) in which we are not interested at present and which is merely possessed but not thematised.

T-Denotation. In order for this imagined entity to be a model of the Sun-Earth system it has to stand for this system, that is, it has to T-refer to this system. So what is the source of T-reference? The answer to this question is that we borrow T-reference from other symbols, for the most part words, which are used in introducing the model. When we look at how the model is explained in textbooks, it is usually either the title of the relevant section or the prose in the introductory paragraph which specifies that what we are dealing with is a model of the Sun-Earth system. This assumes that the reader is familiar with the terms 'Sun' and 'Earth', knows what they refer to, and can make sense of the idea of treating them as a system. In some books also drawings are used to achieve T-denotation. In Young and Freedman's textbook for instance, we see a picture with a big shiny yellow spot in the middle and a little black spot on an ellipse around it (2001, 371), along with a caption explaining that a planet moves around the Sun on an elliptical orbit. In this case it is a mixture of linguistic and pictorial symbols that warrants T-reference. In sum, the T-reference of the Sun-Earth model rests on the reference of the terms or images that are used when the model is introduced.

In passing, let me add an observation confirming my thesis in Chapter 5 that scientific representation cannot be naturalised. Most textbooks do not introduce the above model as a model of the Sun and the Earth in particular, but as one of the Sun and *a planet*, and then mention that this planet could be the Earth, for instance. This is possible because from a mechanical point of view the Earth and the other planets are pretty much the same. So what makes the difference between a model of the Sun-Earth system and the Sun-Venus system, say, is not anything in the model as an object; it is the denotation we stipulate. In one case we say 'this is a model of the Sun-Earth system', in the other case we say 'this is a model of the Sun-Venus system' and so the same entity comes to represent different things – we have two indistinguishable counterparts.[74]

Designation. Aspects and features in the model have to designate aspects and features in the target, which is to say that the aspects of the model have to A-refer to the aspects of the target and that links have to be specified between the features.

Let me deal with A-reference first. As in the case of T-reference, we borrow A-reference from language. When we introduce the model we specify that that the big body in the model stands for the Sun, that the small one stands for the Earth, that the forces in the model stand for the forces in the target, etc. All this is effected by dint of language. We match the aspects in the model and the target by using the respective linguistic expressions to stipulate that one refers to the other. So A-reference, as T-reference, ultimately is borrowed from language. In detail we obtain the following list:

1st The first aspect of the model, the body b_1, stands for the Sun.

2nd The second aspect of the model, the body b_2, stands for the Earth.

3rd The third aspect of the model, its forces, stand for the forces in the target.

4th The fourth aspect of the model, its space-time structure, stands for the space-time structure of the target.

5th The fifth aspect of the model, the motion of the bodies, stands for the motion of the bodies of the target.

This is straightforward. Things get more interesting once we come to the links, because these cannot simply be borrowed from language. What we have to do is to identify for each feature F_i of the model a feature G_i of the target to which it is linked and then figure out, on the basis of all background knowledge available, how the two relate. Roughly, this amounts to the following:

[74] One possible caveat: whether they are really indistinguishable depends how precise the masses are specified. Usually a rough approximation is used, and then the masses of Venus and Earth are roughly the same. If one wants to be exact, then m_2 has to assume the value of the mass of Venus. Then the models are no longer strictly indistinguishable.

1^{st} Aspect: body b_1

Corresponding features in the target:

G_1 = The Sun has mass m_s.

G_2 = Though the surface of the Sun has a lot of local irregularities due to explosions, escaping clouds of gas and other processes going on inside the sun, on a large scale it has roughly spherical shape.

G_3 = The Sun basically is an ensemble of different shells: there is a hot kernel, which is surrounded by different layers of other materials.

Links:

F_1–G_1: By assumption the model has the correct mass built into it. So we have an identity link: $m_s = m_1$

F_2–G_2: When we adopt a certain measure of geometrical similarity the 'almost-spherical shape' of the Sun is close to the shape of b_1. Such a measure of closeness is for instance the following. Assume the centres of the two bodies coincide. Then take the fraction of the volume of the difference between two bodies – i.e. the space that is occupied by one but not the other – over the volume of the ideal sphere. If this fraction is below 0.05, say, then the two are similar. Since the shape of the Sun is almost a sphere and there is nothing in nature that in-principle would prevent it from being an ideal sphere, this link is an ideal limit in the above sense.

F_3–G_3: The Sun consists of layers grouped around a kernel. Moreover it is the case that each of these layers is virtually homogenous; that is, there are only very small density fluctuations within one layer. Hence the mass distribution is virtually spherical. Again, this can be understood as an ideal limit idealisation.

2^{nd} Aspect: body b_2

The specifications of G_4, G_5, and G_6 as well as their links to F_4, F_5, and F_6 are mutatis mutandis the same as the ones for G_1, G_2, and G_3. For this reason there is no need to repeat them here.

3rd Aspect: forces in the system

Corresponding features in the target:

G_7 = The forces acting between Earth and Sun: there are gravitational forces acting between every part of either body and every part of the other. Moreover, there are magnetic interactions due to the fact that both the Sun and the Earth have a magnetic field.

G_8 = All other forces acting on either Sun or Earth: the gravitational forces between either of the bodies and every other piece of mass in the universe and every other magnetic field in the universe.

Links:

F_7–G_7: First observe that the magnetic fields of both Sun and Earth are so weak that their interaction is negligible, at least compared to the strength of the gravitational interaction. For this reason we can neglect them. Next we have to explain how the centre-to-centre interaction in the model relates to the myriad of interactions that take place between every particle in the Sun and every particle in the Earth. At this point we make reference to what is sometimes referred to as Newton's Theorem (Ohanian 1985, 229-31): the gravitational interaction between two spherical mass distributions is the same as though all the mass of each were concentrated at its centre. For this reason, the effective force between the Earth and the Sun is approximately $f=g*m_1*m_2/d^2$, where d is the distance between their centres. The qualification 'approximately' needs to be made because this result rests on four other links – F_2–G_2, F_3–G_3, F_5–G_5, F_6–G_6 – that have been found to be ideal limits rather than identity links.

F_8–G_8: The link between the two is something like a pragmatic simplification. These interactions are small compared to the interaction between the two bodies and since they would make the problem intractable were they taken into account, we just ignore them. But unlike the above links, this link is not harmless. Making the Sun more or less spherical does not change much, or, to put it in other terms, the system's behaviour is structurally stable under slight deformation of the Sun. It is not structurally stable, however, under changes of external forces. It is one of the Poincaré's groundbreaking results that the trajectory of a planet can diverge significantly from its usual ellipse under the influence of external forces. So by adopting this link we have to limit the time span of prediction.

4th Aspect: b_1 and b_2 are located in space-time

Corresponding feature in the target:

G_9 = The space-time background of the Earth and the Sun is relativistic.

Link:

F_9–G_9: Much has been written about the relationship between classical and relativistic space time; but this is not the place to review these debates. What matters at this point is that whatever stance one adopts on this matter, it seems uncontroversial that classical space time is sufficiently close to relativistic space time for the purpose of the Sun-Earth system. Moreover, if a more accurate model is needed, one can replace the classical space-time in the model by a relativistic one without any problem. Nothing else said so far about the model hinges on the space-time background being classical (unlike the forces, for instance, which depend on the links of the shape and the mass distributions) and for this reason it can be replaced without having to change other things as well.

5th Aspect: motion of the bodies

Corresponding feature in the target:

G_{10} = The Sun and the Earth revolve around their common centre of mass.

Link:

F_{10}–G_{10}: This is again a limit type link. It is not true that the Sun is at rest. But the Sun is so much heavier than the Earth (in fact, the Sun's mass is about 750 times the mass of all other planets combined) that the centre of the Sun virtually coincides with the centre of mass of the system as a whole.

This completes my exposition of the Sun-Earth model. Now we are at the point where we could start 'experimenting' with the model and learn more about it. However, in the present case this process does not begin until a mathematised version is available. How mathematics enters the scene is the topic of the next chapter. For this reason I will not say more about the model here and come back to it in Chapter 8, when a mathematical treatment is available.

5. A Note on Two Alternative Accounts of Scientific Representation

In Chapter 1 I noted that the issue of scientific representation has not received much attention so far. There are two exceptions, though, which I have only mentioned in passing. I therefore conclude this chapter by briefly commenting on them and explaining how they relate to the views I have been developing in the last two chapters.

The DDI account of modelling

In a paper entitled 'Models and Representation' (1997) R.I.G. Hughes introduces what he calls the DDI account, which he takes to be a general theory of scientific modelling. On this account, scientific modelling involves three components: denotation, demonstration, and interpretation (hence the acronym 'DDI'). When we present a model of a target system we first posit that this model denotes the target. Then we exploit the fact that models possess an internal dynamics (or an internal set-up of some sort or another) to demonstrate certain conclusions. Finally we have to 'transfer' the conclusions derived in the model back to the target system in order to make predictions, which, according to Hughes, is a step that involves interpretation.

This brief description makes it clear that the DDI account is a diachronic account of how models are used in an investigative process. That is, it explains how we proceed when we exploit a model to gain knowledge about a target system: we first stipulate that the model stands for the target, then prove what we want to know, and finally 'transfer' the results obtained in the model back to the target. Details aside, this picture seems by and large correct.

The problem with the DDI account is that it does not explain *why* and *how* this is possible. Under what conditions is it true that the model denotes the target? What kinds of things are models that they allow for demonstrations? How does interpretation work; that is, how can results obtained in the model be transferred to the target? For short, what are models and how do they relate to the world? These are questions an account of scientific representation has to address, but which are left unanswered by the DDI account. Therefore, if we understand the DDI account as a theory of scientific representation as introduced in Chapter 1, it is seriously defective.

What should make us suspicious, however, is that these questions are not only left unanswered, as a matter of fact at no point in the paper are they even raised. For this reason I think that it is best not to consider the DDI account as an account of scientific representation at all (contrary to what the title of the paper seems to suggest). What the account really provides us with is a view on how models are put to use in a context of investigation. This is an interesting issue, but one that is different from the question of how scientific representation works (as introduced in Chapter 1). For this reason there is no conflict between the DDI account and the views on representation I have been developing so far. They simply tackle different questions.

The inferential conception of scientific representation

The second approach that I consider is Mauricio Suárez' inferential conception of scientific representation (Suárez 2002, 22-32), which he explicitly presents as a general account of scientific representation. Such an account, on Suárez' view, has to meet one major condition: it has to be deflationary. By this he means that an account of scientific representation should not seek to identify a property, call it 'representativity', which all representations possess – just as a theory of truth should not seek to identify a property, truth, which all true sentences possess. Instead, a theory of representation should be restricted to the description of the most general surface features of representation.

Suárez then goes on and provides two conditions – he takes these to be necessary but not sufficient – which, on his view, provide us with the sought-after deflationary theory of representation. These conditions are the following.

> '[inf]: *A* represents *B* only if (i) the representational force of *A* points towards *B*, and (ii) *A* allows competent and informed agents to draw specific inferences regarding *B*.' (*ibid*, 27).

The first condition is designed to assure that *A* and *B* indeed enter into a representational relationship. But on its own, this condition is too weak because also purely conventional stipulations ('let the teacup on my desk represent the universe') can satisfy it. Cases of that sort are ruled out by the second condition, which posits that scientific representations must have cognitive import. That is, we have to be able to learn from the model about the target or, to put it differently, to draw inferences about the target on the basis of the representation (hence the label 'inferential conception of scientific representation').

This raises two questions. First, does a theory of scientific representation really have to be deflationary? Second, does [inf] provide us with such an account? My answers to both questions are negative. Let me discuss them one at a time.

Why we don't want a deflationary account of representation. Whatever the virtues of a deflationary stance in the case of truth, I think that it is a non-starter when it comes to representation. In the first part of his paper, Suárez rightly points out that there is no *physical* property we should be looking for. But from this it does not follow that there is no property at all that we should try to identify. Quite the contrary. To deny that representation can be naturalised (combined with the trivial premise that there are representations) amounts to saying that it is something other than its physical properties that turns an object into a representation, for instance the *use* someone makes of it. This raises the question of what kinds of uses or activities on the side of the user turn an object into a representation. It seems unsatisfactory to say that this capacity of users is irreducible and that no further analysis is either needed or possible (*ibid.*, 25). As I mentioned in Chapter 2, sheer intention to use something in a representational way is not enough to actually make it representational. We do not understand how words, for instance, refer to something beyond themselves merely by saying that a speaker intends words to do exactly that. Of course he does, but this by itself does not answer the question. What we want to know is *how* the speaker achieves this and coming to terms with this puzzle is what theories of reference are supposed to do. So what we have to understand is what exactly a scientist has to do when she wants to use something as a representation of something else. And similar provisos apply when it comes to the user's capacity to draw inferences about the target on the basis of the model. How is this possible? What kind of thing does a model have to be in order to posses cognitive relevance? How do we learn about a part of the world from a model? All these are questions that call for an answer; merely stating that a representation allows us to draw inferences is not enough.

For these reasons I think that [inf] (and with it any other deflationary account one might think of) is just a restatement of the question of how scientific representation works rather than an answer to it. What a theory of representation has to provide us with is an analysis of where representational force comes from and of how it is possible to learn about the world by using a model. But this is to deny that a deflationary account will fit the bill.

However, there is a corollary to the general deflationary attitude as regards representation, which I think remains valid even if we give up deflationism as a general attitude. The fact that a theory of representation has to provide a substantial analysis of what it means for something to represent something else does not imply that there is one, and only one, true story that covers all cases. On the contrary, there are different types of representations, which require a different analysis. This is true within the realm of scientific representation; and it becomes blatantly obvious once we widen the scope of the investigation. Photographs, paintings, words, diagrams, landscapes, and sign posts (Suárez aims to cover all these with his theory) are representations, but we seem to be barking up the wrong tree if we are looking for one particular property that accounts for the representational character of all of them. We may call the acknowledgement of the fact that not all representations work in the same way *methodological nominalism*. This methodological stance is implied by Suárez' deflationism; and I think that it is correct, even if defelationism on the whole is not.

[inf] as a research programme. Naturally, [inf] can be understood as a deflationary account of representation. But as such it suffers from all the defects I have pointed out. However, despite Suárez' insistence on deflationism, [inf] is not necessarily tied to deflationism. At bottom, the problem with [inf] is that its two crucial terms – 'representational force' and 'inference' – are abstract (in the sense specified in Chapter 3) and need fitting out in every particular instance. That is, per se they do not tell us much because they need to be concretised in every individual case. This observation suggests a different reading of [inf]: rather than being an *account* of representation, it is the *general form* of an account of representation. On this reading, [inf] is a blank that has to be filled in every instance of representation. Therefore, what [inf] provides us with is a general *schema* in which we have to discuss the problem of representation. It is an invitation to take different cases of representation – be they paintings or certain scientific models – and ask for each of these, first, where their representational power comes from and, second, what makes them cognitively relevant.[75] Or to put it another way: [inf], rather than being an account of representation, it is something like a *research programme.* Thus understood, [inf] does neither presuppose nor imply that we cannot give a substantive account of how individual cases of representation work.

If I am right about this, the project I have been pursuing so far can be understood as following this research programme. The three-fold account of scientific representation I presented above can be understood as a fitting out of the abstract conditions provided by [inf] for the particular case of scientific models. As a consequence, there is no tension between the two accounts. Rather, they are complementary.

There is an important heuristic consequence to be drawn from my reading of [inf]. The way in which Suárez talks about [inf] – he repeatedly emphasises that [inf] only provides us with necessary but not sufficient conditions for representation – suggests that in order for an analysis of representation to be complete something over and above [inf] is needed. Then it is our task to figure out what that is. On my reading this is not true. Our task is not to find anything over and above [inf] simply because there is no such thing. [inf] by itself is the most general form of a theory and

[75] This seems to square with Suárez' intuitions when, at one point, he characterises [inf] as 'a scheme that will be filled in differently in each instance of representation' (2002, 27).

nothing needs to be added to it. Accordingly, the conditions provided by [inf] are both sufficient and necessary rather than only necessary. What we have to do, however, is to fit out the abstract terms that figure in it. That is, we have to explain what 'representational force' amounts to in each particular case and we have to say what features of the particular representation at hand allow it to function cognitively and in what way. Once we have done that, our task is completed.

Chapter 8

The Use of Mathematics in Scientific Modelling

1. Introduction: Models and the Applicability of Mathematics

In Chapter 5 I argued that equations are distinct from models. Nevertheless, equations (and other mathematical expressions) play an important role in scientific modelling and many scientists spend most of their time investigating mathematical items of some sort. This raises two questions. First, how is the use of mathematics in scientific modelling possible at all? Second, how does the use of mathematics in scientific modelling square with the views developed in the previous chapters? These are the two questions I address in this chapter. As it turns out, an answer to the second question is implied in the answer to the first. When we account for how mathematics can be used, we at once understand how it can be integrated into the overall picture of representation I have been canvassing in the last three chapters.

For this reason, let me begin by saying a few more words about the first question. When asking how the use of mathematics is possible in the context of scientific modelling we are asking an old question in new guise. The old quandary is how mathematics is applicable to the non-mathematical world. Wigner famously remarked that 'the enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious and that there is no explanation for it' (Wigner 1960, 2).[76] One need not go as far as seeing the applicability of mathematics as an inexplicable miracle, but the question remains: how is the use of mathematics in science possible? Or to put it another way: how does mathematics hook onto the world?

[76] Many other physicists have made similar remarks; see Steiner (1998) for a compilation of quotes.

This question is a time-honoured philosophical puzzle – it can be traced at least to Plato's Timaeus – and there is a considerable body of literature on the subject. But given that the present chapter belongs to the constructive part of the thesis, I do not start with a review and a critique of the different positions. Rather I begin by developing a positive account and defer a discussion of alternative views to the end of the chapter, where I briefly comment on Platonist, formalist, intuitionist, logicist and descriptivist takes on the problem of applicability.

My answer to the problem of the applicability of mathematics is structuralist. Given my critical stance toward structuralism so far this may seem odd, or even incoherent. It is not. I only argued that structuralism does not fit the bill as a full-fledged theory of representation, which does not imply a denial of the fact that structures do play an important role in theoretical science. What structuralism really gives us is a theory about the applicability of mathematics, and not a theory of scientific representation. In what follows I present an explicit formulation of a structuralist view of the applicability of mathematics and show how it ties in with my other views on scientific representation.

This gives rise to the following plan. In the next section I first outline and develop a structuralist view on the applicability of mathematics. In Section 3 I continue the study of the Sun-Earth system, introduced in Chapter 7, and show how mathematics is applied to the world in this case. In the last section I briefly discuss alternative accounts and explain why I think that they do not fare as well as structuralism.

2. A Structuralist Conception of the Applicability of Mathematics

The overall picture

The leading idea of a structuralist conception of the applicability of mathematics is that mathematics applies to the world because it describes structures, and these structures can be 'embodied' in a physical system. On this view, then, a law expressed in mathematical terms is construed as saying that a certain structure is instantiated in a particular part of physical reality, or that some material objects are arranged in such a way that they 'fill' the positions of a structure described by mathematics. This granted, the properties of the structure – being invariant under certain transformations, for instance – are also properties of the object that instantiates the structure. In other words, what we know about the structure itself can be translated into knowledge about the object possessing the structure.[77]

This account of the applicability of mathematics rests on three premises (Ketland 2001, 42):

Premise 1: Pure mathematics is the study of structures (structuralism about pure mathematics).

Premise 2: Structures can be instantiated in physical objects.[78]

Premise 3: Applying mathematics involves 'translating' knowledge about a structure into knowledge about a concrete physical system.

As a simple example consider arithmetic. On a structuralist view, arithmetic is the study of the natural number structure, the structure of any infinite sequence of objects that has an initial element and a successor relation. This structure (or rather one of its initial segments) can be instantiated in a string of teacups on a table, say, simply by stipulating that one is the initial element and then lining up the others in a way that imposes a follower relation on them. We then can translate theorems of arithmetic (for instance that 2+3=5) into knowledge about the cups because they instantiate (the initial segment of) the structure.

In order to render the structuralist account of the application of mathematics benign we have to flesh out and justify these premises. What does that involve? Premise 1 is a claim falling within the domain of the philosophy of pure mathematics, where it has been widely discussed (at least) since the groundbreaking work of Frege and Dedekind. In what follows I will provide a statement of the structuralist position and add some remarks by way of motivation; but it is beyond

[77] This view has been articulated in different ways by Shapiro (1983, 1997, 2000), Resnik (1997) and Hellman (1989).

[78] A terminological remark: Ketland, Shapiro and Resnik for the most part use the verb 'exemplify' when they talk about a system having (or instantiating) a structure. I avoid this terminology because 'exemplification' is used in a specific sense in the literature on representation (Goodman 1968, Ch. 2; Elgin 1983, Ch. 5; 1996, Ch. 6), which is not intended at this point.

the scope of this chapter to review the debate about structuralism in mathematics and comment on it. The main contention of Premise 2 has been extensively discussed in Chapters 3 and 5. So we can build on the results previously obtained. As far as Premise 3 is concerned, it seems that there is not much one can say about it from a general philosophical perspective. Rather, the problem of translation is one that belongs to science and if some difficulties that may come up in one case or another are philosophical in nature, they belong to the philosophy of certain special sciences such as the philosophy of space and time.

Before getting into a discussion of these premises, let me detail the overall picture of the use of mathematics in scientific modelling that I suggest. Schematically, what we have is the following:

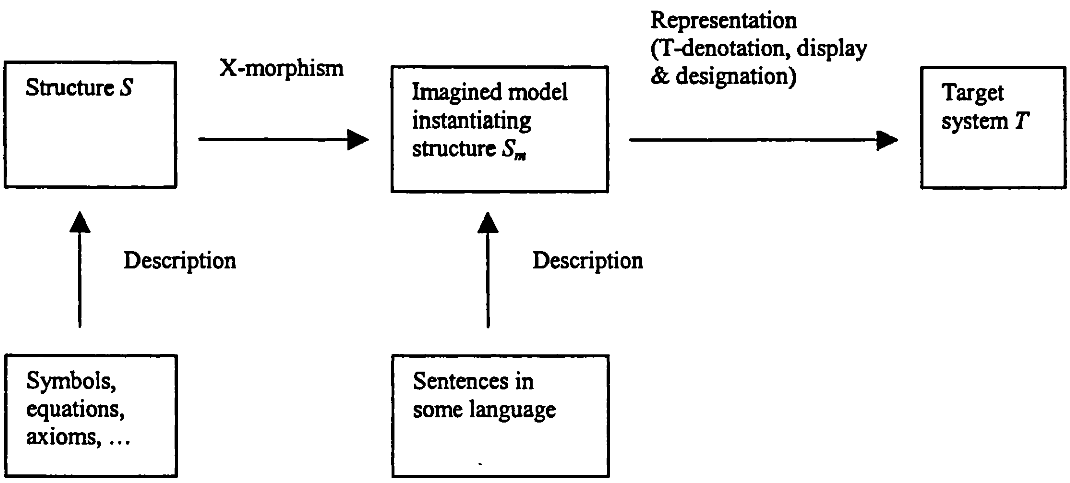

Fig. 7: The structuralist conception of the applicability of mathematics.

The relation between the imagined model and the target system is the representation relation specified in the previous chapter. The description relation between sentences and the imagined model has been explicated in Chapter 5. The other description relation, the one between mathematical expressions (axioms, equations, etc.) and structures is warranted by Premise 1. The fact that the imagined model is taken to instantiate a structure is backed by Premise 2. These are the unproblematic parts of the scheme (that is, of the scheme *as a scheme* – the premises themselves are not unproblematic at all). What is in need of further elucidation is the introduction of two structures – S and S_m – and the X-morphism that is postulated to connect the two. 'X-morphism' is a placeholder for any mapping that can connects two structures such as an isomorphism, a homomorphism, or an embedding.[79] In the simplest case, the mapping connecting S and S_m is an isomorphism. In this case, the introduction of two separate structures is actually superfluous and one could remove S from the schema and take the mathematical description to refer directly to the instantiated structure S_m. But things are not always that simple. In many cases the structure S that is described by a calculus is 'bigger' than the structure the model instantiates. For instance, many structures studied by mathematicians and used in the sciences have an infinite domain, while there may only be a finite number of parts available in the model. In cases like this, the structures S and S_m are not identical and the mapping between the two is not an isomorphism, but an embedding or a homomorphism.[80] Or the structure may be 'too big' in other ways. In electronic engineering, for instance, one often uses complex numbers to represent the voltage of an electric current. But only the real part of the complex number corresponds to a physical quantity; the imaginary part has been added for mathematical convenience.[81] In what sense a structure can be 'too big' and what kind of mappings can connect it to the instantiated structure is a question that will receive different answers in different contexts and there is little one can say about it in general. What matters is that the overall picture one adopts is such that it makes room for various possibilities, which is the case with the above scheme.

Premise 1: structuralism in the philosophy of mathematics

What is the subject matter of mathematics? On a Platonist view, mathematics deals with mathematical objects, numbers for instance, which are conceived along the lines of physical objects in that they are taken to have ontic and epistemic independence from one another. The existence of the number three, say, is no more dependent on

---

[79] The fact that a certain X-morphism holds between S and S_m is commonly referred to as a 'representation theorem'.

[80] Cases of that sort are described in Redhead (2001). He calls the parts of S that do not correspond to anything in S_m 'surplus structure'.

[81] More elaborate examples of that sort include the S-matrix approach to scattering and gauge theories (see Redhead 2001).

the existence of other numbers than the existence of one chair is on the existence of other chairs; and we can know about the number three in isolation from other numbers. Mathematics, then, is the science about these objects.

Structuralists vigorously reject this 'object view' of mathematics. On a structuralist outlook, the subject matter of mathematics is not individual mathematical objects but rather the structure in which these objects are arranged. The objects themselves are of no importance. They are construed as 'featureless, abstract positions in structures' (Resnik 1997, 4). The essence of a natural number or a space-time point, for instance, is their relations to other natural numbers or space-time points and they do not have any 'internal constitution'. Their identity is fixed uniquely with respect to their relationships to other numbers or points and there is strictly nothing to them over and above this; mathematical objects, as far as they exist, are nothing but a place, or position, in a structure. For this reason, mathematics is not a science about mathematical objects but about abstract structures, where structures are understood to be unspecific structures in the sense introduced in Chapter 3. Or to put it another way: mathematics is the study of structures as such.[82]

To repeat, it is beyond the scope of this chapter to argue for this position and I endorse it without further discussion. Those who are not convinced by the structuralist take on mathematics may be reconciled with the view on the *application* of mathematics that I am presenting in this chapter by the following observation. The structuralist about the *application* of mathematics can do with less than the structuralist about *pure* mathematics. The latter claims that mathematics is about patterns *and nothing else*; the former does not need to impose such a restriction. In order to get a structuralist view on the *application* of mathematics to work one only has to grant that mathematics is at least *inter alia* about structures, regardless of whether it also describes mathematical objects, mental constructs, or what have you. For instance, a Platonist who believes in the reality of mathematical object can still hold that certain structures are instantiated by both abstract mathematical things and

[82] Dedekind is commonly credited for being the founding-father of the structuralist camp in the philosophy of mathematics. Succinct contemporary statements of the view can be found in Resnik (1997), Shapiro (1997, 2000) and Hellman (1989, 1996, 2001).

concrete objects in the world and that this is how these two otherwise separated realms relate.[83, 84]

Let me close this short comment on structuralism in mathematics by noting that there are different brands of structuralism, corresponding to different ontological attitudes towards structures. Structural Platonists like Michael Resnik (1997) and Stewart Shapiro (1983, 1997, 2000) take structures to be 'ante rem' universals. On this view, structures exist independently of the physical systems instantiating them. It is then reasonable to speak of the natural number structure or of the structure of a Euclidean space regardless of whether there exist any systems in the physical world that possess these structures.[85] More empiricist-minded philosophers are reluctant to accept Platonic entities of that sort. In this vein, Geoffrey Hellman (1989, 1996) has suggested that we only accept concretely instantiated structures as real and in all other cases substitute assertions of logical possibility for mathematical existence claims. On this view, structures with an infinite domain, for instance, do exist but only in the sense that their existence is logically possible, where logical possibility is accounted for by the use of a primitive modal operator for second order logical possibility.[86] These differences have important consequences for the technical formulation of how the structure S, which is described by a certain calculus, relates to the structure S_m, which is instantiated by a physical system. The modal structuralist faces the problem of connecting the possibilities introduced at the level of pure mathematics with the actual world and it turns out that this requires the

[83] According to Shapiro (1983, 545) even a nominalist like Hartry Field can adopt such a point of view.

[84] Some care is needed, however, when it comes to translating mathematical results into knowledge about a physical system. Only results whose proof relies on nothing but structural features can be carried over to the physical realm. If a Platonist, for instance, uses properties of mathematical objects that go beyond structural features (if this is possible) in a proof, the result thus proven cannot be transferred to the empirical world if the connection between the two realms is only structural.

[85] For the sake of completeness let me note that there is a second brand of structuralism one might label as 'Platonist': set theoretic structuralism (see Hellman 2001 for an exposition of this position and a comparison with the other two options). As far as the application of mathematics is concerned, 'ante rem' and set theoretic structuralism are on equal footing and for this reason I will not say more about them here.

[86] For a discussion of the internal problems this view faces see Resnik (1997, 67-80).

introduction of further modal operators. The structural Platonist, on the other hand, can posit mappings between the two domains in a rather straightforward way. However, on either account – and this is the salient point – one connects a mathematical structure to a structure instantiated in the physical system.

Premise 2: structures and physical objects

Premise 2 posits, to repeat, that structures can be instantiated in objects. This claim has been discussed at length in Chapters 3 and 5 and so it suffices at this point to refer to the results previously obtained: a target system has a structure *as represented* in a model.

Premise 3: translation

Premise 3 states that applying mathematics involves 'translating' knowledge about a structure into knowledge about a concrete physical system. It is trivial both *that* this is what we want – after all, this 'export of knowledge' is the objective for using mathematics in the sciences in the first place – and *that* it is actually possible – it happens all the time when we successfully use mathematics in a model. The case study in the next section provides examples for both claims. The question is *how* this translation takes place. The translation of mathematical knowledge into knowledge about the model depends on several factors: the nature of the particular structures involved, the scientific problem at hand, and the X-morphism chosen (if S_m is embedded in S, the way in which facts about S translate into facts about S_m differs from how they translate if the two structures enter into a homomorphism, for instance). From a purely philosophical point of view, however, there is not much one can say about how this translation takes place. One simply has to work through each case in its own terms; and therefore the problem of translation is one that has to be discussed within the particular sciences and the philosophies of these, rather than within a general theory of representation. In some cases, such as arithmetic, a translation of mathematical facts is straightforward. Once we establish a one-to-one mapping from the tea cups on the table into some initial segment of the natural number structure, arithmetical statements readily translate into statements about tea cups. In other cases results may not be forthcoming so easily. How, for instance, do facts about symplectic manyfolds translate into facts about a mechanical system is not prima facie clear; and in yet other cases it is still an unsolved problem how exactly this translation has to be effected (quantum mechanics is one case in point). But regardless of whether the case at hand is problematic or straightforward, to figure how this translation works is a problem that has to be addressed within the specific scientific discipline at stake – or a foundational philosophical discipline thereof – and not within a general theory of representation.

Summing up

Before illustrating these claims in a case study, let me summarise the picture of the applicability of mathematics that I suggest and let me add some comments by way of clarification. The above schema needs to be improved in one essential respect. As it stands, it does not account for the insight gained in Chapter 3 that not every structure faces reality directly; many structures are only related to other structures and not to the physical world. However, this 'cascade' has to come a halt and the 'lowest' structure has to be grounded in reality by means very different from those that are used to connect one structure to another structure. This insight can easily be built into the above schema. To this end, let S_1, ...S_p be structures (where p is an arbitrary integer) and let 'TDD' be the acronym for 'T-denotation, display, and designation'. Then we obtain the following picture.

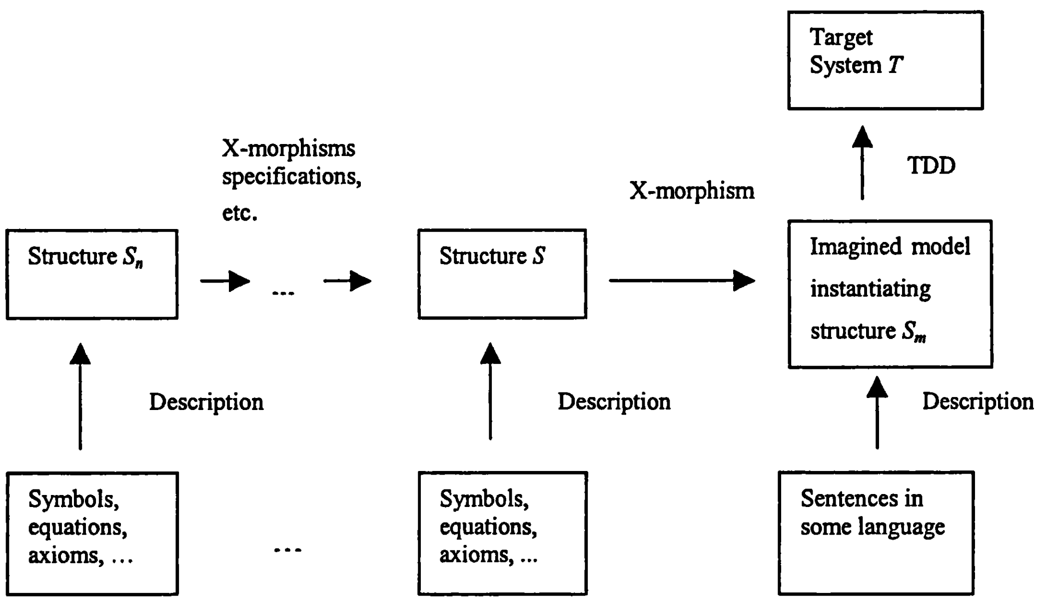

Fig. 8: The augmented structuralist conception of the applicability of mathematics.

Looking at this diagram one might now be puzzled and ask: where in all that is *the* model and which one of these numerous flashes stands for *the* representation relation? From what I have said so far it is clear that my answer to this question is that *the* model is the imagined model and *the* representation relation is the flash labelled 'TDD'; and I am inclined to regard the other structures as part of a theory and the relations holding between them as intra-theoretical relations of some sort. Admittedly, there is the air of convention to this choice. Adherents of the semantic view will insist, I suspect, that all structures S_i are models and that every relation between two of these structures is a representation relation. As long as there is agreement on the overall picture of how the different parts of the scheme integrate, this is a merely verbal matter and there is no harm in calling structures 'models' and X-morphisms 'representation relations' (after all, this is how logicians use the term 'representation'). But this prima facie harmless choice of terminology turns into a highway to fiasco when we lose the grand picture, take a part of it – two structures and an isomorphism, say – for the whole and then claim that all there is to scientific representation are structures and isomorphisms. Structures at some point need to latch onto the physical world and at that point we are faced with problems and questions that are entirely different from the ones we encounter when relating two structures; and it is at this point that we need the conceptual tools developed in Chapters 5 through 7.

3. Illustration: The Mathematical Treatment of the Sun-Earth System

In this section I illustrate the above claims by discussing how the Sun-Earth system, introduced in Chapter 7, is treated mathematically. More specifically, the aim of this section to show how the three premises above bear out in the case at hand, how the structure of the model is 'parasitic' upon an imagined entity and how not all scientifically relevant structures face reality directly.

The Structure of Classical Particle Mechanics

As I mentioned in Chapter 7, the present model of the Sun-Earth system is a classical model; that is, the bodies are located in classical space-time and their motion is supposed to follow the laws of classical mechanics. From a philosophical point of view, nothing hinges on this choice. One could replace the classical space-time background by a relativistic one and treat their motion relativistically. The philosophical lessons about how mathematics is applied would be unaltered. So I stick to classical mechanics mainly for the sake of convenience.

Before a structural treatment of the Sun-Earth system can be offered, we need to analyse the structures of classical particle mechanics (PCM). Such an analysis has been presented by Balzer, Moulines and Sneed in their treatise on the structuralist programme. In the sequel I by and large follow their treatment (*ibid.*, 29-34, 103-108, 180-191; see also Muller 1998, 259-66). However, certain facilitations are inevitable because a complete reconstruction of CPM is neither needed nor possible in this chapter and not all aspects of their reconstruction of CPM matter for my purposes.

The leading idea is to understand CPM as a theory dealing with all possible motions of particles, which are considered as mass-points in space where it is assumed that these motions are caused by forces (*ibid.*, 29). For this reason, the structure of CPM must roughly be the following pentuple: (particles, time, space, mass, force), where the number of particles is non-zero but finite, time is associated with the real numbers R, space is associated with $R \times R \times R$ (R^3), mass is associated with a positive real number (R^+), and force is associated with a real 3-vector.

Following this basic idea, one can define a potential CPM – I will come to the qualification 'potential' below – as $\Omega=[P, (T, \tau), (S, \sigma), m, f, r, R, N]$, where P, T, S, m, f, τ, σ, r are such that they satisfy certain requirements (*ibid.*).[87] What are these

[87] Two remarks with regard to the notation are needed. First, my notation differs slightly from the one used by Balzer, Moulines and Sneed; I hope that my changes render the presentation more intuitive. Second, the current discussion of the structure of CPM differs slightly from the one used in Chapter 2 in that I here do not first list all the objects, then all the relations, and finally all the operations. This would be inconvenient because the structure of CPM naturally divides into several substructures (space, time, etc.), and it would be contrived (if not confusing) to 'decompose' them into their individuals, relations and operations and list all of them separately. For this reason, Ω contains a list of

requirements? Before presenting precise definitions, let me state the intuitive ideas in an informal way:

- P is the set of particles, where particles are understood as point masses.
- T is time, construed as *physical* time. It is supplied with the usual classical structure, the real number structure, by the mapping τ, which maps T onto the real numbers R.
- S is space, construed as *physical* space. It is supplied with the usual classical structure, the one of R^3, by the mapping σ, which maps S onto R^3.
- m is the function assigning a mass to every particle.
- f is the force function. Naturally, f depends on the particles and time (the dependence on space comes in via the dependence on particles, which are located in space). Furthermore, there is usually more than one force in the system. To do justice to that, we introduce an index i in the force function to distinguish between different forces and obtain: $f(p, t, i)$. For instance, i=1 can be the gravitational force acting on a certain particle; i=2 can be the electrostatic force acting on the same particle, and so on. The forces acting on particle p can be added by standard vector addition: $\Sigma_i f(p, t, i)$.
- r is the function assigning each particle at each time a position in space.

Bearing this in mind, we can state the following formal version of these requirements (*ibid.*, 30):

- P, T, S are non-empty sets; P is finite.
- τ: $T \to R$ and σ: $S \to R^3$ are bijective mappings.
- r: $P \times T \to S$.
- σ, r, and τ are such that $\sigma \circ r_p \circ \tau^{-1}$: $R \to R^3$ is smooth for all $p \in P$, where τ^{-1} is the inverse of τ, r_p is the mapping we obtain from r by holding the argument p fixed, and '$\circ$'is the multiplication of two functions.
- m: $P \to R^+$.
- f: $P \times T \times N \to R^3$.

substructures rather than individuals etc. But this is merely a matter of presentation and does not affect the essence of the concept of a structure.

Finally, why is Ω the structure of a 'potential' CPM? The point is that no law of motion has been specified as yet and therefore it is not clear how forces give raise to motion. In order to obtain CPM we have to add Newton's second law in a structural version. To facilitate notation, define $\rho_p:=\sigma^\circ r_p{}^\circ\tau^{-1}$ and let D be a differential operator. Then $\Omega=[P, (T, \tau), (S, \sigma), m, f, r, R, N]$ is a CPM iff its constituents have the properties specified above, and

$$m(p)*D^2\rho_p(\theta)= \Sigma_i f(p, \tau^{-1}(\theta), i) \qquad \text{(Newton's law)}$$

holds true for all $p \in P$ and $\theta \in R$, where θ is a 'mathematical' instant of time, i.e. the image of a physical instant of time $t \in T$ under the mapping τ, and '*' denotes multiplication (as in Chapter 7).

Note that only structural 'ingredients' (sets of things along with relations in which they enter and operations acting on them) have been used so far. In particular, at no point have we made use of a formal language to describe the structure. However, this becomes unavoidable once we want to learn about the structure and prove theorems. We then choose particular co-ordinates – Cartesian ones, for instance – to co-ordinatise space and time and certain units for mass and force. We then can write down Newton's equation of motion:

$$m*\mathrm{d}^2\underline{x}/\mathrm{d}t^2 = f(\underline{x}, t), \text{ where } \underline{x}=(x, y, z),$$

which is true of Ω by construction. But, and this is the important point, we could also choose any other set of co-ordinates (spherical, cylindrical, elliptical, hyperbolic, or what have you). In these co-ordinates the equation looks different, but it is still true of Ω because it is merely a different way of describing the same thing, namely the structure of CPM.

The structure of the Sun-Earth System

Now let us apply CPM to the Sun-Earth system. To this end we have to specify what the particles are and what forces act between them. And here we already get into trouble. The Sun and the Earth both consist of a myriad of mass particles and between each pair of particles there is a gravitational attraction. Taking this at face value, one would have to put all the particles in the set P and then write down all the forces $f(p, \tau^{-1}(\theta), i)$, where p would range over all particles of both the Sun and the Earth. Little physical intuition is needed to see that such a heroic attempt – which would require solving something like 10^{28} coupled differential equations – would not take us anywhere. But fortunately we can do better. As I mentioned in Chapter 7, Newton's theorem tells us that the gravitational interaction between two bodies with spherical mass distributions is the same as if all the mass of each were concentrated at its centre. This rings a bell. In Chapter 7 we modelled the Sun-Earth system as consisting of two objects of exactly that sort (needless to say, we did so with this theorem in mind!). So we can forget about all the particles in the two bodies and just treat them as two point masses, located at the centre of the spheres. This said, we define the relevant parts of the structure as follows:

- Let p_s stand for the point particle associated with the Sun, p_e for the one associated with the Earth. Then we have $P=\{p_s, p_e\}$.
- Let m_s and m_e stand for the mass of the Sun and the Earth, respectively. Then the mapping m is as follows: $m(p_s)=m_s$ and $m(p_e)=m_e$.
- The force acting between p_s and p_e is gravitation. If we let $\underline{e}_{es}$ stand for the unit vector pointing from the Earth to the Sun we have: $f(p_e)=\underline{e}_{es}*g*m_s*m_e/d^2$ and $f(p_s)=-\underline{e}_{es}*g*m_s*m_e/d^2$, where d is the distance between the centre of the Sun and the centre of the Earth. The other arguments of the force function – time and i – can be dropped in this case because there is only one force present (by assumption, see Chapter 7) and gravitation is not time dependent.

The other elements of Ω (space, time, etc.) do not need to be defined again since they are not in any way dependent on the particular physical system we are dealing with. Gathering the pieces together we immediately obtain the structure of the Sun-Earth system:

$$\Omega_{se}=[\{p_s, p_e\}, (\mathrm{T}, \tau), (\mathrm{S}, \sigma), m(p_s)=m_s \text{ and } m(p_e)=m_e,$$
$$f(p_e)=\underline{e}_{es}*g*m_s*m_e/d^2 \text{ and } f(p_s)=-\underline{e}_{es}*g*m_s*m_e/d^2, r, R, N],$$

where the elements in Ω_{se} satisfy Newton's law.

We now can choose co-ordinates and write down the equation of motion. A clever choice in this case is polar co-ordinates because these are 'adapted' to the rotational symmetry of the system. Given this, we can start proving theorems about the system, such as that angular momentum is conserved, that energy is conserved, that the Laplace-Runge-Lenz vector is a conserved quantity, that p_e moves on an elliptical orbit with p_s at one focus (Kepler's first law), that the radial line segment from p_s to p_e sweeps out equal areas in equal time (Kepler's second law), and that the square of the period is proportional to the cube of the semi major axis of the planetary orbit (Kepler's third law). How to do this is not straightforward, but there is no need here to show how it can be done (treatments of this problem can be found in any advanced mechanics textbook). What matters in the context at hand is that all these theorems are theorems about the structure Ω_{se}.

This illustrates the three premises of the structuralist account of the applicability of mathematics. First, Newton's equation of motion describes a structure, namely Ω in the general case and Ω_{se} if details of the Sun-Earth system are plugged in. This is true by construction, so there is nothing more to show here. Second, the Sun-Earth system has a structure, namely Ω_{se}. But it has this structure only – and this is the salient point – *as represented* in the imagined model introduced in Chapter 7. If we do not assume that the planets have spherical mass distribution, that the only force present in the system is gravitation and that they are placed in a classical space time background, the system does not have the structure Ω_{se}. It will still have structures, for sure, but other and often more complicated ones. Third, we can translate mathematical knowledge first into knowledge about the imagined model and then about the target system. For instance, that angular momentum is conserved implies that the motion of p_s and p_e takes place within a plane. Since the space-time links from model to target are almost identity, this straightforwardly translates into the fact that the motion of the Earth virtually takes place in a plane. For the same reasons the theorems about the trajectories of p_s and p_e (Kepler's three laws) translate into features of the motion of the Earth as well: the Earth moves on an ellipse, and so on. The fact that energy is conserved translates into the fact that the Earth will keep revolving around the Sun and does not come closer and closer to it and ultimately collapse into it. Other features of the mathematical structure, however, have no intuitive physical translation. For instance, that the Laplace-Runge-Lenz vector is a conserved quantity does not translate into anything observable about the system. But there is no harm in that. First, Premise 3 does not say that *all* mathematical knowledge is translatable, it only says that some is. Second that we don't yet know how to translate this matter of fact does not imply that there is no translation; it may well be that one day we find out how to do it.

Further lessons

The above example illustrates further aspects of the structuralist account of the applicability of mathematics.

First, the example of CPM makes it plain that not every structure faces reality directly: Ω does not. But some structures do, in this case Ω_{se}. And the way in which two structures are related is totally different from how a structure and the target system are related. The relation between Ω and Ω_{se} is specification. The relation between Ω_{se} and the target system is more convoluted: Ω_{se} is possessed by the imagined entity described in Chapter 7; this entity in turn represents the target – the Sun-Earth system – via T-denotation, display, and designation; and for this reason the target possesses Ω_{se} *as represented* in this model.

Enough has been said by now about how the structure Ω_{se} relates to the world, so let me add some qualifications on the concept of specification, but without going into too much detail (for details see Balzer, Moulines and Sneed 1987, Ch. 4). A theory does not deal only with one particular structure, but with a whole family of structures. This family is not merely a more or less random collection of things; rather it has a particular 'architecture'. Balzer, Moulines and Sneed call this architecture a 'theory net'. The leading idea is that there are different elements of a theory, which enter in a kind of hierarchical order according to how basic they are. At the top we have a general principle (or several of them). This principle is then specified by applying it to particular situations. From a model-theoretic point of view, this amounts to adding more special laws to the already existing fundamental laws and thereby 'carving out' of the set of actual models a subset determined by these more severe restrictions (*ibid.*, 169). In CPM, for instance, we have Ω at the top. We then can impose several restrictions on Ω, which happens basically by imposing restrictions on the admissible force functions. For instance, we can stipulate that the force cannot be explicitly time dependent, that it has to be symmetrical under certain transformations, and so on. In this fashion we can work our way down in the hierarchy until we reach structures incorporating very concrete force laws such as Hooke's law. Each specification provides us with a so-called theory element. If we represent this hierarchy of theory elements graphically, we obtain an array that has the structure of an inverted tree. But each tree has a bottom end at which we find structures that cannot be further specified. These have to be connected to reality in the way described.

Second, the example with the Sun-Earth system nicely illustrates my remark in Chapter 3 that the choice of certain structure is not a 'one-way enterprise'. We do not first carve the system, make some assumptions and then see what structure we get. We often chose a very general 'background structure' to begin with – in the present case CPM – we want to put to use and much of what we then do on the concrete level is guided, or at least motivated, by this abstract structure. If we did not know CPM and that Newton's theorem allows us to reduce a myriad of interactions to a single one, we would not necessarily model the Sun and the Earth as ideal spheres with spherical mass distributions. In short, the final result is often determined from both the bottom and the top.

Third, it is an interesting observation that advanced textbooks (such as Goldstein 1980, Scheck 1992, or Landau and Lifschitz 1984) do not discuss the modelling assumptions outlined in Chapter 7 at all. They immediately start with the mathematical treatment of the problem and don't spill any ink on things like linking. Does that show that I have overemphasised the importance of substantial models and that, after all, mathematics by itself is enough? I think that is wrong for the following reason. These authors evade talk about these things not because they do not play any role, but simply because there is no need to: their readers already know. By the time one can read a book like Goldstein's one has done at least one or two years of basic university physics, and in the books one reads in these years there is a lot about the relevant modelling assumptions! The sources I used in the last chapter are good examples. These are first year textbooks that every physics student has to read and by the time she reads Goldstein, the whole story about ideal spheres and so on is so familiar that there is no need to repeat it. So the absence of material modelling in advanced mechanics textbooks does not play into the hands of the radical structuralist (i.e. one who claims that models are nothing but structures).

The same thing happens when new disciplines are established. Hofbauer and Sigmund afford us with a telling example. In the preface to the second edition of their book on evolutionary games they write: 'In our former book, it took us 150 pages of biological motivation to tentatively introduce the notion of a replicator equation. This is no longer warranted today: replicator dynamics is a firmly established subject [...] and our old volume definitely looks dated today.' (1998, xi). This boils down to saying that now people know what the biology 'on the ground' is which warrants the use of these mathematical tools. There is no point in repeating it all over again. But this does not mean that the concrete biology in the background does not play any role, it is just not necessary to mention it every time. Or to put it another way, the substantial models are implicit in the mathematical specifications.

4. A Remark on Alternative Points of View

In this chapter I developed a structuralist view of how mathematics applies without saying much by way of motivation, let alone justification of this approach. Such a justification would involve a discussion and critique of the known alternatives as well as an argument to the conclusion that structuralism fares better than each of them. Clearly, this is beyond the scope of this chapter. But in order not to leave the issue up in the air altogether, I will now add some brief remarks indicating what I take these advantages to be and why it seems that a structuralist account of the applicability of mathematics fares better than its alternatives.

Let me begin with the advantages. Though my discussion of structuralism in Part II was mainly critical, there is a positive flipside to it in that it also indicates how structures can be brought back into the picture (a fact that I exploited above). So structures, though dismissed as representational tools, never got entirely out of sight. For this reason, going for a structuralist view of the application of mathematics seems to be the most natural move. But over and above convenience, structuralism also enjoys the advantage over possible alternatives that it allows us to restore a certain continuity in the philosophical discussion about the nature of scientific theories in that the use of structures makes it possible to connect the current approach to much of the work done on this issue during the last four decades. If I am right on the question of how scientific representation works, many of the claims about scientific theorising put forward in the context of the semantic view need revision, but they need not be thrown overboard altogether.

As far as possible alternatives are concerned, there are good reasons to believe that structuralism is the best game in town. My reasons for believing so are basically the ones put forward by Shapiro (1983). In brief, these are the following.

(1) *Formalism.* On a formalist view, mathematics consists of no more than the manipulation of characters according to rules. If this is correct, the systematic correspondence between certain mathematical theorems and facts in the world remains a mystery. If all a mathematician does is fiddling around with meaningless symbols, it is not clear why mathematics should have more of a relationship to the world than any other rule guided activity such as playing chess or dancing a ballet.[88] There is simply no reason to assume that proving a theorem, say, should shed light on anything beyond the rules.

(2) *Logicism.* The basic slogan of logicism is that mathematics is logic. In one version – 'translation logicism' – this is understood as saying that mathematical statements, like sentences of logic, are true or false solely in virtue of their form. A mathematical theorem, then, is a truth of logic. This renders the applicability of mathematics incomprehensible because it clashes with the fact that the description of an interesting empirical phenomenon is never a truth of logic. On another reading – 'postulate logicism' – mathematics is the study of logical consequences of uninterpreted axioms. Thus construed, mathematics per se does not have a subject matter at all because its terms are not endowed with reference. On this view, then, the role of mathematics in science is to uncover logical connections between certain statements. The problem with this view is that it faces the same difficulties as descriptivism: it can only account for applications involving branches of mathematics that have a straightforward interpretation of their basic terms, which is the case only in particular circumstances.

[88] A detailed critique of formalism can also be found in Ketland (2001).

(3) *Platonism.* Platonists hold that mathematics is the study of mathematical objects that populate a transempirical, non-mental realm. How then do the entities in this 'mathematical world', whose existence is independent of the physical world, relate to empirical objects? The independence of these two realms does not, of course, preclude that there is a relationship between them, but Platonists have not come up yet with an explanation of what this relationship is that would allow us to account for the sophisticated use modern science makes of mathematics.

(4) *Intuitionism.* Intuitionists agree with Platonists that mathematics is the study of mathematical objects, but they take these to have no existence outside the human mind. Mathematical objects are mental constructs. Within this framework one might try to account for the applicability of mathematics by postulating a relationship between the material world and the portions of the human brain that do the mathematics. This programme, however, is in need of articulation and it remains to be seen whether it could work out.

(5) *Descriptivism.* Finally, one of the first answers that might come to mind when confronted with the question of how mathematics is applied in the empirical sciences is that mathematical expressions can be reinterpreted such that they come to refer to things in the world. On that view, mathematics is a language of sorts and it is applicable in the context of the empirical sciences because it can be used to describe objects in the world.

Despite its initial plausibility and intuitive appeal this view is untenable for at least three reasons. First, descriptivism can only account for the application of branches of mathematics that have straightforward interpretations of their languages in terms of physical things. But this is not always the case. There is nothing in the physical world in terms of which the languages of abstract algebra or complex operators, for instance, could be interpreted; but, as is well known, they are very useful in the context of physical theories that do shed light on some parts of reality. Many mathematical theories require a prior ontology that by far outstrips what both physical theory and common sense acknowledge. Most mathematical theories, including the ones invoked in the sciences, require an uncountably infinite domain and it is at least an open question whether we can find that many objects in the physical world. So if we take mathematics at face value, there are just not enough things in the world for it to be plausible that the language of a mathematical theory can be interpreted in terms of physical objects.[89]

Second, the arguments marshalled against descriptivism about models in Chapter 5 apply mutatis mutandis to the present case as well. As plain descriptions most successful applications of mathematics are false. The surface of a real table is not a Euclidean plane, the orbit of the planets are not ellipses, the motion of a pendulum is not a sine function, and so on. As in the case of models, we seem obliged to admit that the description really is a description of some object other than the actual target system and that this object bears on the target indirectly by standing in a certain relationship to it. So mathematical expressions are not direct descriptions of the target.

Third, as Ketland (2001, 18-19) points out, the claim that mathematical symbols refer to concrete physical entities when mathematics is applied in the sciences is descriptively wrong. Often, mathematical symbols refer to mathematical objects, even when used in a scientific context. The symbol '$A_\mu(x)$' for instance, when used in electromagnetic theory, refers to a function assigning to each physical space-time point a quadrupel of real numbers; and it does not refer to any 'physical stuff'.

For these reasons the applicability of mathematics cannot be explained in terms of reinterpretation of mathematical symbols.

[89] This point has been made, in different contexts though, by both Shapiro (1983, 531) and Resnik (1997, 204-5).

Glimpses Beyond

In the eight chapters of this thesis I first framed what I take to be the problem of scientific representation, then argued that none of the currently available accounts of scientific theorising has a tenable answer to offer and finally suggested a positive view of how scientific representation works. In this last part I would like to briefly discuss some of the things that have been left out and mention some questions that naturally arise from the views that I have been putting forward.

The variety of representational strategies. In Chapter 7 I suggested an account that acknowledges linking as one of the constitutive relationships of scientific representation. But the question of what linking is has only received a partial answer. As I indicated, the general considerations concerning the character of linking need to be complemented by an identification and discussion of different modes of representation, which will finally result in something like a dictionary of representational strategies. This is a task I could not possibly undertake within the confines of the present thesis. But the question is one that needs to be addressed: what representational strategies are there and how exactly do they work?

Systematising the zoo of models. The literature on models has been growing quickly over the last four decades, and with it the number of different types of models that philosophers recognise. A cursory survey yields the following (by no means exhaustive) list: phenomenological models, probing models, developmental models, symbolic models, impoverished models, testing 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. As it stands, this abundance is quite bewildering and it calls for systematisation. The theory of representation as developed in Part III of this thesis provides us with a starting point to achieve such a systematisation once we realise that many of these labels (if not all of them) make either implicit or explicit reference to the way in which a model represents its target, or to the way in which it fails to do so. A first systematisation can be achieved by looking at how the three relations that I take to be constitutive of scientific representation do or do not obtain in each of these cases. A probing model, for instance, is one that fails on the counts of T-denotation and designation and only displays certain properties. Models of this sort can be employed to find out something about properties themselves, for instance how they interact or whether they are compatible at all. A didactic model is one that fares poorly on the side of display in the sense that it only displays a very limited range of well-known properties and once built, it is not able to reveal anything new. Examples for models of this kind are the familiar ball-and-stick chemistry models of a molecule. They are useful to students, but not to scientists because the balls and the sticks do not possess any properties that could be exploited for further research. Caricature models, finally, are models that have weak links either because the properties the model displays are very different from the ones the target is taken to possess, or because we simply do not know what the link between them really is. Needless to say, these characterisations are rough and ready and are in need of qualification. I merely mention them to indicate in what way the conception of representation suggested in Chapter 7 could be put to use in such a classifying enterprise. This systematisation can be further refined once a list of representational strategies is available. It is then rather straightforward to say that every linking strategy gives rise to a type of model: ideal limit models, analogue models based on shared properties, and so on.

The epistemology of imagined models. How do we learn about the world from a model that only exists in our imagination? Material models can more or less be understood along the lines of common experimental method, but what about imagined models? What constraints are there to the construction of imagined models? How do we obtain 'results' in such a model and what is their status? In what way can they be brought to bear on the physical world? These and other questions need to be addressed to understand how imagined models work. There has recently been some interest in questions similar to these in connection with thought experiments (see Reiss 2003 and references therein) and a discussion of imagined models can certainly draw on these debates. But a detailed account of how we gain knowledge from imaginary models is still much needed.

Is structural realism a blind alley? In the early 20th Century, Henri Poincaré (1902) and Bertrand Russell (1927) independently of each other – and for different reasons – put forward the view that our knowledge of the external world is solely structural and that science only describes structure. On this view, all we can know is structure and we have to remain completely agnostic as regards all the rest. This position, now commonly referred to as *structural realism*, has generated great interest among philosophers of science and it has recently been advocated in different guise by many, most notably by John Worrall (1989) and Elie Zahar (2001). Steven French (1998) and James Ladyman (1998) gave the doctrine a metaphysical twist and formulated what they call *ontic structural realism*. On this view, we are just fooled when we believe that a structure is a structure *of something else*; what we have to realise is that ultimately all there is to nature is structure; that is, all that exists is structure.

The argument put forward in Chapter 3 for the conclusion that structures are abstract casts doubt on both the epistemic and the ontic version of structural realism. In a nutshell, the point is that if the instantiation of a structure rests on more concrete, non-structural facts, it seems that we have to know about these facts in order to know about the structure. Knowledge about structures then is derivative in that it rests on knowledge about more concrete features of the world. But this is incompatible with epistemic structural realism. And similarly for the ontic version: if a system possesses a structure only relative to some more concrete features, it just cannot be that structure is all that exists. It seems that this is a promising line of argument against structural realism, but one in need of qualification. There are many different versions of structuralism and many different ways of justifying the position and it needs to be shown in detail for each of these versions how the argument from abstractness undercuts its success.

Fictionalism. Knowledge is commonly taken to be justified, true belief. Despite much disagreement about what each of these three defining components amounts to, the basic idea is widely accepted. So one would expect that the knowledge we gain from models conforms to this picture. This, however, does not seem to be the case because there is a tension between the requirement of truth and the admission of non-identity links between the features of a model and its target. A model only yields a true picture of the system when the properties it displays coincide with the ones the system possesses. But this is not normally the case. Models employ all kinds of idealisations and simplifications and for this reason do not provide us with a true picture of their target system. In this sense models are fictions. So we are faced with the problem of how we gain knowledge about the world from fictions. How can the use of fictional entities illuminate anything beyond the fiction? What, if anything, does the use of frictionless planes and ideal rational agents reveal about real planes and real agents? This is the problem of fictionalism.

As Arthur Fine observes in a programmatic essay (1993), the last systematic discussion of this problem within the standard philosophy of science literature was Israel Schffler's in 1963. Although this may be a slight overstatement,[90] the general thrust of this remark seems to be correct: a systematic discussion of fictionalism is still needed.

The most natural way to go seems to relax the conditions for knowledge. In this vein Catherine Elgin (1997) suggests that what we should aim at is not truth but understanding. This suggestion has a great deal of prima facie plausibility, but it leaves us with the problem of what understanding is. Another line of argument, one that is closer to the view on representation developed above, would also sacrifice truth and replace it by one of the (yet to be specified) strategies of representation. This would lead to notions like 'idealising knowledge'. It is not clear, however, what this sort of knowledge would be knowledge of and in what way it would illuminate aspects of the real world.

It is not clear what the right response to the problem of fictionalism is. What is clear, however, is that a response is needed and that the issue deserves more attention than it has received so far.

Scientific representation versus other kinds of representation. Throughout this thesis, the analogy between object-to-object representation in the visual arts (especially in painting) and in the sciences was of great heuristic value. Important parts of my discussion of scientific representation explicitly draw on parallel issues

[90] Goodman and Elgin's plea for a reconception of philosophy (1983, Ch. 10; also Elgin 1997) and Cartwright's notion of physics as theatre (1983, Ch. 7) point in this direction (although neither of these authors uses the label 'fictionalism'). Moreover, there is a considerable body of literature on idealisation and approximation, which in some sense also can be understood as contributing to a discussion of fictionalism.

in the philosophy of art. This raises the question of how far this analogy stretches. What are the commonalties of scientific and artistic representation and in what sense do they differ? They serve a different purpose (though cognitivists would dispute this), are created in a different way, and are put to different uses. That much is obvious. But in what 'technical' aspects do they differ? What are the means, the 'tools', used in scientific representation that are not used in an artistic context and vice versa? Dispite pioneering work by Goodman (1968) and Elgin (1997), much remains to be said about this issue.

Bibliography

Achinstein, Peter (1968): *Concepts of Science. A Philosophical Analysis*. Baltimore.

Balzer, Wolfgang, C. Ulises Moulines and Joseph D. Sneed (1987): *An Architectonic for Science. The Structuralist Program*. Dordrecht.

Beardsley, Monroe C. (1981): *Aesthetics. Problems in the Philosophy of Criticism*. 2nd ed., Indianapolis and Cambridge.

Bell, John and Moshé Machover (1977): *A course in Mathematical Logic*. Amsterdam and Oxford.

Black, Max (1960): 'Models and Archetypes', in: Max Black: *Models and Metaphors. Studies in Language and Philosophy*. Ithaca and New York 1962, 219-43.

– (1970): 'How do Pictures Represent?', in: Ernst H. Gombrich, Julian Hochberg, and Max Black: *Art, Perception, and Reality*. Baltimore and London, 95-130.

Bogen, James and James Woodward (1988): 'Saving the Phenomena', *Philosophical Review* 97, 303-52.

Boolos, George S. and Richard C. Jeffrey (1989): *Computability and Logic*. 3rd ed., Cambridge.

Bourbaki, Nicolas (1957): *Eléments de Mathémathique XXII. Livre I: Théorie des Ensembles*. Paris.

Braithwaite, Richard B. (1953): *Scientific Explanation*. Cambridge.

Bueno, Otavio (1997): 'Empirical Adequacy: A Partial Structure Approach', *Studies in the History and Philosophy of Science* 28, 585-610.

– (1999): 'What is Structural Empiricism? Scientific Change in an Empiricist Setting', Erkenntnis 50, 59-85.

Carnap, Rudolph (1938): 'Foundations of Logic and Mathematics', in: Otto Neurath, Charles Morris and Rudlph Carnap (eds.): *International Encyclopaedia of Unified Science. Vol. 1*, Chacago, 139-213.

Carroll, Noël (1999): *Philosophy of Art. A Contemporary Introduction*. London and New York.

Cartwright, Nancy (1983): *How the Laws of Physics Lie*. Oxford.

– (1989): Nature's Capacities and their Measurement. Oxford.

– (1999): *The Dappled World. A Study of the Boundaries of Science.* Cambridge.

– and Henry Mendell (1984): 'What Makes Physics' Objecs Abstract?', in: James Cushing, C.F. Delaney, and G.M. Gutting (eds.): *Science and Reality.* Notre Dame, 134-52.

Chakravartty, Anjan (2001): 'The Semantic or Model-Theoretic view of Theories and Scientific Realism', Synthese 127, 325-45.

Da Costa, Newton C. and Steven French (1990): 'The Model-Theoretic Approach to the philosophy of Science', *Philosophy of Science* 57, 248-65.

– (2000): 'Models, Theories, and Structures: Thirty Years On', Philosophy of Science 67, Proceedings, S116-127.

Danto, Arthur (1981): *The Transfiguration of the Commonplace. A Philosophy of Art.* Cambridge/Ma and London.

Denis, Maurice (1909): 'From Gauguin and van Gogh to Neo-Classicism', reprinted in: Charles Harrison and Paul Wood (eds.): *Art in Theory 1900-2000.* Oxford, Maldon a.o. 2003, 46-51.

Devitt, Michael and Kim Sterelny (1987): *Language and Reality.* Cambridge/MA.

Duhem, Pierre (1906): *La Théorie Physique, son Objet et sa Structure.* 2nd ed., Paris 1914, unaltered reprint Paris 1997. Engl. trans. by Philip P. Wiener: *The aim and Structure of Physical Theory.* Princeton 1954.

Dummett, Michael (1991): *Frege. Philosophy of Mathematics.* London.

Elgin, Catherine Z. (1983): *With Reference to Reference.* Indianapolis.

– (1996): Considerate Judgement. Princeton.

Feynman, Richard (1963): *The Feyman Lectures in Physics. Volume I.* Ed. by Richard Feynman, Robert B. Leighton and Mathew L. Sands. Commemorative Issue, Redwood City and Reading (Ma.) 1989.

Fine, Arthur (1993): 'Fictionalism', *Midwest Studies in Philosophy* 18, 1-18.

French, Steven (1998): 'On the Withering Away of Physical Objects', in: E. Castellani (ed.): *Interpreting Bodies: Classical and Quantum Objects in Modern Physics.* Princeton, 93-113.

– (1999): 'Models and Mathematics in Physics: The Role of Group Theory', in: Jeremy Butterfield and Constantin Pagonis: *From Physics to Philosophy.* Cambridge, 187-207.

– (2000): 'The reasonable effectiveness of Mathematics: Partial Structures and the Application of Group Theory to Physics', *Synthese* 125, 103-20.

– (2002): 'A Model-Theoretic Account of Representation', unpublished Manuscript, presented at the PSA 2002 in Milwaukee.

– and James Ladyman (1999): 'Reinflating the Semantic Approach', *International Studies in the Philosophy of Science* 13, 103-121.

Frigg, Roman (2002): 'Self-Organised Criticality - What it is and what it Isn't', *Measurement in Physics and Economics Discussion Paper Series*, DP MEAS 19/02. Forthcoming in *Studies in History and Philosophy of Science*.

Giere, Ronald N. (1985): 'Constructive Realism', in: Paul M. Churchland and Clifford A. Hooker (eds.): *Images of Science. Essays on Realism and Empiricism with a Reply from Bas C. van Fraassen*. Chicago and London, 75-98.

– (1988): *Explaining Science*. Chicago.

– (1999): 'Visual Models', in. Ronald Giere: *Science Without Laws*. Chicago, 118-146.

– (2002): 'How Models Are Used to Represent Physical Reality', unpublished Manuscript, presented at the PSA 2002 in Milwaukee.

Goldstein, Herbert (1980): *Classical Mechanics*. 2nd ed. Reading (Ma.) and London.

Goodman, Nelson: (1968): Languages of Art. 2nd ed. Indianapolis and Cambridge 1976.

– (1972): 'Seven Strictures on Similarity', in: Nelson Goodman: *Problems and Projects*. Indianapolis and New York, 437-446.

– and Catherine Z. Elgin (1978): *Reconceptions in Philosophy and Other Arts and Sciences*. Indianapolis and Cambridge.

Graham, Gordon (2000): *Philosophy of the Arts. An Introduction to Aesthetics*. 2nd ed. London and New York.

Hellman, Geoffrey (1989): *Mathematics without Numbers: Towards a Modal-Structural Interpretation*. Oxford.

– (1996): 'Structuralism Without Structures', *Philosophia Mathematica* 3, 100-123.

– (2001): 'Three Varieties of Mathematical Structuralism', *Philosophia Mathematica* 9, 184-211.

Hempel, Carl G. (1965): 'Aspects of Scientific Explanation', in: Carl G. Hempel: *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York, 331-496.

Hesse, Mary (1963): *Models and Analogies in Science*. London and New York.

Hodges, Wilfrid (1997): *A Shorter Model Theory*. Cambridge.

Hofbauer, Joseph and Karl Sigmund (1998): *Evolutionary Games and Population Dynamics*. Cambridge.

Hughes, R. I. G. (1997): 'Models and Representation', in: Philosophy of Science 64, Proceedings, S325-36.

Ketland, Jeffrey (2000): 'Newman's Theorem and the Relativity of Exemplification', unpublished manuscript.

– (2001): 'How the Laws of Mathematics Apply', unpublished manuscript.

Kitcher, Philip (1993): *The Advancement of Science. Science Without Legend, Objectivity without Illusions*. Oxford a.o.

Krantz, D.H., R.D. Luce, P. Suppes and A. Tversky (1971): *Foundations of Measurement. Volume I, Additive and Polynomial Representations*. New York and London.

Kroes, Peter (1989): 'Structural Analogies between Physical Systems', *British Journal for the Philosophy of Science* 40, 145-54.

Ladyman, James (1998): 'What is Structural Realism?' *Studies in the History and Philosophy of Science* 29, 109-24.

Landau, Lev D. and E. M. Lifschitz (1984): *Lehrbuch der Theoretischen Physik. Vol. 1: Mechanik.* 11th ed. Berlin.

Layman, Ronald (1987): *Using Scott Domains to Explicate the Notions of Approximation and Idealised Data.* Philosophy of Science 54, 194-221

– (1991): 'Thought Experiments by Stevin, Mach and Gouy: Thought Experiments as Ideal Limits and Semantic Domains', in: Tamara Horowitz and Gerald J. Massey (eds.): *Thought Experiments in Science and Philosophy.* Savage/MD, 167-91.

Lyas, Colin (1997): Aesthetics. London.

Machover, Moshé (1996): *Set Theory, Logic and Their Limitations.* Cambridge.

McAllister, James W. (1997): 'Phenomena and Patterns in Data Sets', *Erkenntnis* 47, 217-228.

Morgan, Marry S. and Margaret Morrison (eds.) (1999): *Models as Mediators. Perspectives on Natural and Social Science.* Cambridge.

– (1999): 'Models as Mediating Instruments' in: Marry S. Morgan and Margaret Morrison 1999, 10-37.

Morgan, Mary S. (1999a): 'Learning from Models', in: Marry S. Morgan and Margaret Morrison 1999, 347-88.

– (1999b): 'Models, Stories, and the Economic World', *Research Memoranda in History and Methodology of Economics*. University of Amsterdam.

Morgan, Mary S. and Marcel Boumans (1998): 'The Secrets Hidden by Two-Dimensionality: Modelling the Economy as a Hydraulic System', *Research Memoranda in History and Methodology of Economics.* University of Amsterdam.

Morrison, Margaret (1999): 'Models as Autonomous Agents', in: Marry S. Morgan and Margaret Morrison 1999, 38-65.

Moulines, C. Ulises *et al.* (2002): *Structuralism.* Special issue of *Synthese* 130 (1).

Muller, Frederik A. (1998): *Structures for Everyone*. Amsterdam.

Mundy, Brent (1986): 'On the General Theory of Meaningful Representation', *Synthese* 67, 391-437.

Nagel, Ernest (1961): *The Structure of Science. Problems in the Logic of Scientific Explanation*. New York 1961; repr. Indianapolis/Cambridge 1979.

Newman, M. H. A. (1928): 'Mr. Russell's "Causal Theory of Perception"', *Mind* 37, 137-148.

Ohanian, Hans C. (1985): *Physics*. 2nd ed., New York and London.

Peetz, Dieter (1987): 'Some Current Philosophical Theories of Pictorial Representation', *British Journal of Aesthetics* 27, 227-37.

Pitkänen, Risto (1981): 'On the Analysis of Pictorial Representation', *Acta Philosophica Fennica*, 31, Vol. 4.

Poincaré, Henri (1902): *La Science et l'Hypothèse*. Paris. Translation: *Science and Hypothesis*, New York 1905.

Psillos, Stathis (1995): 'The Cognitive Interplay between Theories and Models: The Case of 19th Century Physics', in: William E. Herfel, Wladiyslaw Krajewski, Ilkka Niiniluoto and Ryszard Wojcicki (eds.): *Theories and Models in Scientific Process*. Poznan Studies in the Philosophy of Science and the Humanities 44, 105-33.

– (1999): *Scientific Realism. How Science Tracks Truth*. London.

Putnam, Hilary (1981): *Reason, Truth, and History*. Cambridge.

Redhead. Michael (2001): 'The Intelligibility of the Universe', in: Anthony O'Hear (ed.): *Philosophy of the New Millennium*. Supplement to *Philosophy*, Royal Institute of Philosophy Supplement 48, 2001, 73-90.

Reiss, Julian (2003): 'Causal Inference in the Abstract or Seven Myths about Thought Experiments', *Causality: Metaphysics and Methods Research Project Technical Report* 03/02. LSE.

Rickart, Charles E. (1995): *Structuralism and Structure. A Mathematical Perspective*. Singapore a.o.

Rohrlich, Fritz (1989): 'The Logic of Reduction: The Case of Gravitation,' in: *Foundations of Physics* 19, 1151-70.

Russell, Bertrand (1919): *Introduction to Mathematical Philosophy*. London and New York 1993.

– (1927): *The Analysis of Matter*. London 1992.

Scheck, Florian (1992): *Mechanik. Von den Newtonschen Gesetzen zum Deterministischen Chaos*. 3rd ed. Berlin and Heidelberg.

Schier, Flint (1986): *Deeper into Pictures. An Essay on Pictorial Representation*. Cambridge.

Sellars, Wilfrid (1965): 'Scientific Realism or Irenic Instrumentalism', in: Robert S. Cohen and Marx W. Wartofsky (eds.): *Boston Studies in the Philosophy of Science. Volume Two: In Honor of Philipp Frank*. New York 1965, 171-204.

Shapiro, Stewart (1983): 'Mathematics and Reality', *Philosophy of Science* 50, 523-48.

– (1991): *Foundations without Foundationalism. A case for Second-Order Logic*. Oxford.

– (1997): *Philosophy of Mathematics. Structure and Ontology*. Oxford.

– (2000): *Thinking About Mathematics*. Oxford.

Short, T.L. (1988): 'Hypostatic Abstraction in Empirical Science', *Grazer Philosophische Studien* 32, 51-68.

Steiner, Mark (1998): *The Applicability of Mathematics as a Philosophical Problem*. Cambridge/MA and London.

Suárez, Mauricio (1999): 'Theories, Models, and Representations', in: L. Magnani, N. J.Nersessian and P. Thagart (eds.): *Model-Based Reasoning in Scientific Discovery*. New York, 75-83.

– (2002): 'The Pragmatics of Scientific Representation', *Centre for Philosophy of Natural and Social Science Discussion Paper Series* DP66/02.

Suppe, Frederick (ed.) (1977): *The Structure of Scientific Theories*. Urbana and Chicago.

– (1989): *The Semantic Conception of Theories and Scientific Realism*. Chicago.

Suppes, Patrick (1960a): 'A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences', in: Patrick Suppes: *Studies in the Methodology and Foundations of Science. Selected Papers from 1951 to 1969*. Dordrecht 1969, 10-23.

– (1960b): 'Models of Data', in: Patrick Suppes: *Studies in the Methodology and Foundations of Science. Selected Papers from 1951 to 1969*. Dordrecht 1969, 24-35.

– (1967): 'What is a Scientific Theory', in: Sidney Morgenbesser (ed.): *Philosophy of Science Today*. New York, 66-67.

– (1970): *Set-Theoretical Structures in Science*. Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford.

– (1988): 'Representation Theory and the Analysis of Structure', *Philosophia Naturalis* 25, 254-68.

Swoyer, Chris (1991): 'Structural Representation and Surrogative Reasoning', *Synthese* 87, 449-508.

Teller, Paul (2001): 'Whither Constructive Empiricism', *Philosophical Studies* 106, 123-150.

van Fraassen, Bas C. (1980): *The Scientific Image*. Oxford.

– (1981): 'Theory Construction and Experiment: An Empiricist View', Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 2, 663-677.

– (1989): *Laws and Symmetry*. Oxford.

– (1991): *Quantum Mechanics*. Oxford.

– (1992): 'From Vicious Circle to Infinite Regress, and Back Again', Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 2, 6-19.

– (1994): 'Interpretation of Science; Science as Interpretation', in: Jan Hilgevoord (ed.): *Physics and our view of the World.* Cambridge, 169-187.

– (1995) 'A Philosophical Approach to Foundations of Science', *Foundations of Science* 1, 5-9.

– (1997): 'Structure and Perspective: Philosophical Perplexity and Paradox', in: M. L. Dalla Chiara et al. (eds.): *Logic and Scientific Methods.* Dordrecht, 511-30.

– (2001): 'Structure: It's Shadow and Substance.' Forthcoming in: P. Ehrlich and R. Jones (eds.): *Reverberations of the Shaky Game: Essays in Honor of Arthur Fine.* Chicago 2001.

– (2002a): *The Empirical Stance.* New Haven and London.

– (2002b): 'Scientific Representation: Flouting the Criteria', unpublished manuscript, presented at the PSA 2002 in Milwaukee.

Wartofsky, Marx W. (1979): *Models: Representation and the Scientific Understanding.* Dordrecht.

Wigner, Eugene (1960): 'The Unreasonable Effectiveness of Mathematics in the Natural Sciences', *Communications on Pure and Applied Mathematics* 13, 1-14.

Wollheim, Richard (1987): *Painting as an Art.* London.

Woodward James (1989): 'Data and Phenomena', *Synthese* 79, 393-472.

Worrall, John (1989): 'Structural Realism: The Best of Both Worlds?', *Dialectica* 43 (1989), 99-124.

Young, Hugh D. and Roger Freedman (2000): *University Physics. With Modern Physics.* 10th ed., San Francisco and Reading (MA).

Zahar, Elie (2001): *Poincaré's Philosophy. From Conventionalism to Phenomenology.* Chicago and La Salle.

Zeman, J. Jay (1982): 'Peirce on Abstraction', *Monist* 65, 211-229.