

Fiction and Scientific Representation

Roman Frigg¹

In: Roman Frigg and Matthew Hunter (ed.): *Beyond Mimesis and Nominalism: Representation in Art and Science*, Berlin and New York: Springer, 2010, 97-138.

1. Introduction

Scientific discourse is rife with passages that appear to be ordinary descriptions of systems of interest in a particular discipline. Equally, the pages of textbooks and journals are filled with discussions of the properties and the behaviour of those systems. Students of mechanics investigate at length the dynamical properties of a system consisting of two or three spinning spheres with homogenous mass distributions gravitationally interacting only with each other. Population biologists study the evolution of one species procreating at a constant rate in an isolated ecosystem. And when studying the exchange of goods, economists consider a situation in which there are only two goods, two perfectly rational agents, no restrictions on available information, no transaction costs, no money, and dealings are done immediately. Their surface structure notwithstanding, no competent scientist would mistake descriptions of such systems as descriptions of an *actual* system: we know very well that there are no such systems. These descriptions are descriptions of a *model-system*, and scientists use model-systems to *represent* parts or aspects of the world they are interested in. Following common practice, I refer to those parts or aspects as *target-systems*.

What are we to make of this? Is discourse about such models merely a picturesque and ultimately dispensable *façon de parler*? This was the view of some early twentieth century philosophers. Duhem (1906) famously guarded against confusing model building with scientific theorising and argued that model building has no real place in science, beyond a minor heuristic role. The aim of science was, instead, to construct theories with theories understood as classificatory or representative structures systematically presented and

¹ To contact the author write to r.p.frigg@lse.ac.uk. For further information visit www.romanfrigg.org.

formulated in precise symbolic language. With some modifications this view also became dominant among the logical positivists of the Vienna Circle and the Berlin Group; see, for instance, Carnap (1938) and Hempel (1965).

Early resistance against this understanding of science came from Campbell (1920) and Hesse (1963), who emphasised the importance of models to scientific theorising. The tides changed in the 1970s and 1980s. On the one hand the positivist view that theories were partially interpreted logical calculi (now referred to as the ‘syntactic view of theories’) was replaced by the so-called semantic view of theories, according to which a theory simply *is* a collection of models; see Suppe (1977). Parallel, but by and large unrelated to the rise of the semantic view, a tradition of philosophy of science arose that emphasises the importance of scientific practice to philosophical analysis, and so places models again at the heart of a philosophical account of science; see the essays collected in Morgan and Morrison (1999). Hence, current philosophies of science of all stripes agree with a characterisation of science as an activity aiming at representing parts of the world with the aid of scientific models.

For this reason the questions of what scientific models are and how they represent have become central to the concerns of philosophers of science. This chapter proposes a novel approach to the issue of models and representation, one that draws essentially on the analogy between models and literary fiction. But before we can sketch the outlines of this account, some setting up is needed.

As the above examples show, when presenting a model scientists offer us the description of a hypothetical system, one that does not actually exist in nature, which they proffer as an object of study.² Scientists sometimes express this fact by saying that they talk about ‘model-land’; see for instance Smith (2007, 135). The rationale for doing so is that this hypothetical system has two desirable properties. First, it is chosen such that it is easier to study than the target-system and therefore allows us to derive results. Second, it is assumed to represent its target system, and representation is something like a ‘licence to draw inferences’. Representation allows us to ‘carry over’ results obtained in the model to the target-system and hence it enables us to learn something about that system by studying the model.

² Some scientific models are material objects (for instance the wood models of cars that we put into a wind tunnel), but most models are not of this kind. I here focus on models that are, in Hacking’s (1983, 216) words, ‘something you hold in your head rather than your hands’.

Thus, scientists actually perform two acts when they propose a model: they introduce a hypothetical system as the object of study, and they claim that this system is a representation of a target-system of interest. This is reflected in the promiscuous usage of the term ‘model’ in the sciences. On the one hand ‘model’ is often used to denote the hypothetical system we study (e.g. when we say that the model consists of two spheres). On the other hand it is employed to indicate that a certain system represents, or stands for, another system (e.g. when we observe that the Newtonian model of the solar system misrepresents its target in various ways). In practice, however, these two acts are often carried out in tandem and scientists therefore rarely, if ever, clearly distinguish the two.

While this may well be a legitimate way of proceeding efficiently in the heat of battle, it is detrimental to philosophical analysis where it is germane that these two acts be kept separate. In this chapter I endeavour to clearly separate these two acts and to present an analysis of each. To this end, let me first introduce some terminology. I use the term ‘model-system’ to denote the hypothetical system proffered as an object of study. I call those descriptions that are used to introduce the model-system as ‘model-descriptions’. Representation then is the relation between a model-system and its target-system. The term ‘model’ could refer to either the model-system or representation, or the combination of the two, or yet other things; I will therefore avoid it in what follows. I use the term ‘modelling’ to refer to the practice of devising, describing and using a model-system. In this more regimented language, the two acts performed in utterances of the kind mentioned above are, first, presenting a model-system and specifying some of its essential properties, and, second, endowing this model-system with representational power.

This separation may do some violence to common sense, which regards representational power as an intrinsic to things that are models and sees this dissociation of model-systems from representation as artificial at best. Common sense is wrong. It has been pointed out variously – and in my view correctly – that, in principle, anything can be a representation of anything else.³ Representations are not a distinctive *ontological* category and it is wrong to believe that some objects are, *intrinsically*, representations and other are not. It is one question to ask what an object is in itself; but it is quite a different one to ask what, if anything, an

³ The point is Goodman’s (1976); in recent years Teller (2001), Giere (2004) and Callender and Cohen (2006) have discussed it with special focus on scientific representation.

object represents and in what way. Taking model-systems to be intrinsically representational is a fundamental mistake. Model-systems, first and foremost are objects of sorts, which can, and *de facto* often are, used as representations of a target-system. But the intrinsic nature of a model-system does not depend on whether or not it is so used: representation is extrinsic to the medium doing the representing.

Hence, understanding scientific modelling can be divided into two sub-projects: analysing what model-systems are, and understanding how they are used to represent something beyond themselves. The first is a prerequisite for the second: we can only start analysing how representation works once we understand the intrinsic character of the vehicle that does the representing. Coming to terms with this issue is the project of the first half of this chapter. My central contention is that models are akin to places and characters of literary fictions, and that therefore theories of fiction play an essential role in explaining the nature of model-systems. This sets the agenda. Section 2 provides a statement of this view, which I label the *fiction view of model-systems*, and argues for its *prima facie* plausibility. Section 3 presents a defence of this view against its main rival, the structuralist conception of models. In Section 4 I develop an account of model-systems as imagined objects on the basis of the so-called pretence theory of fiction. This theory needs to be discussed in great detail for two reasons. First, developing an acceptable account of imagined objects is mandatory to make the fiction view acceptable, and I will show that the pretence theory has the resources to achieve this goal. Second, the term ‘representation’ is ambiguous; in fact, there are two very different relations that are commonly called ‘representation’ and a conflation between the two is the root of some of the problems that (allegedly) beset scientific representation. Pretence theory provides us with the conceptual resources to articulate these two different forms of representation, which I call p-representation and t-representation respectively. Putting these elements together provides us with a coherent overall picture of scientific modelling, which I develop in Section 5.

While p-representation turns out to be internal to pretence theory (and hence is explained by pretence theory itself), an analysis of t-representation has to draw on different resources. This resource is maps. In Section 6 I present an analysis of how maps represent their target systems and claim that the general structure of this account doubles as the general structure of t-representation. In other words, the view that I am proposing is that one can think of the model-system as a kind of a ‘generalised map’ and explain how it represents (t-represents) its

target along the lines of how maps represent their targets. In Section 7 I use this view to analyse the Newtonian model of the solar system and show that it not only gives a plausible understanding of what happens in this model, but even makes important features of it visible that are usually concealed. Far from being an idle philosophical pastime, the fiction view of models, I claim in conclusion, can actually help us to better understand what is involved in the representational activities essential to scientific models.

2. Model-Systems and Fiction

What kind of things are model-systems? Referring to them as ‘model-systems’ has a homely ring to it which obscures the fact that we don’t know what they are. As we have seen, the descriptions in question are not descriptions of any actual system. So what, if anything, are they descriptions of? What sense can we make of the common practice to qualify claims about such systems as true or false? And how do we find out about the truth and falsity of such claims?

My answers to these questions take as their starting point the realisation that model-systems share important aspects in common with literary fiction. This is more than just an interesting but eventually inconsequential observation. My claim is that thinking about model-systems as being akin to characters and places in literary fiction provides essential clues to solving pressing problems in the philosophy of science. In other words, drawing an analogy between scientific modelling and literary fiction is not idle musing; it is the driving force behind an approach to scientific modelling that aims to provide an understanding of a central aspect of scientific practice.

The core of the fiction view of model-systems is the claim that model-systems are akin to places and characters in literary fiction. When modelling the solar system as consisting of ten perfectly spherical spinning tops physicists describe (and *take themselves* to be describing) an imaginary physical system; when considering an ecosystem with only one species biologist describe an imaginary population; and when investigating an economy without money and transaction costs economists describe an imaginary economy. These imaginary scenarios are tellingly like the places and characters in works of fiction like *Madame Bovary* and *Sherlock Holmes*. These are scenarios we can talk about and make claims about, yet they don’t exist.

Although hardly at the centre of attention, the parallels between certain aspects of science and literary fiction have not gone unnoticed. It has been mentioned by Maxwell, and occupied centre stage in Vaihinger's (1911) philosophy of the 'as if'. In more recent years, the parallel has also been drawn specifically between models and fiction. Cartwright observes that 'a model is a work of fiction' (1983, 153) and later suggests an analysis of models as fables (1999, Ch. 2). McCloskey (1990) regards economists as 'tellers of stories and makers of poems'. Fine notes that modelling natural phenomena in every area of science involves fictions in Vaihinger's sense (1993, 16), and Sklar highlights that describing system "as if" they were systems of some other kind is a royal route to success (2000, 71). Elgin (1996, Ch. 6) argues that science shares important epistemic practices with artistic fiction. Hartmann (1999) and Morgan (2001) emphasise that stories and narratives play an important role in models, and Morgan (2004) stresses the importance of imagination in model building. Sugden (2000) points out that economic models describe 'counterfactual worlds' constructed by the modeller. I have defended the view that models are imaginary objects in my (2003) and my (2009), and Grüne-Yanoff and Schweinzer (2008) emphasise the importance of stories in the application of game theory.⁴ Moreover, Godfrey-Smith (2006) has recently set out what amounts to the most explicit and forceful statement of the fiction view of model-systems now available.

What we have to recognise, though, is that the analogy between model-systems and fiction is only a starting point. If put forward without further qualifications, explaining model-systems in terms of fictional characters amounts to explaining the unclear by the obscure. In fact, fictional entities are beset with philosophical problems that are so severe that avoiding fictional entities altogether would appear to be a better strategy. Fictional entities do not exist: there is no woman called Emma Bovary and there is no detective Sherlock Holmes. Yet they have some kind of reality: we think about them, we talk about them, and they are objects of our emotions. Fictional entities are the subject matter of discussions, and claims about them can be true or false: we say that it is true that Holmes is a detective but false that he is a ballet dancer. How can this be if there is no Holmes? And how can sentences containing the name 'Holmes' even be meaningful if Holmes does not exist? It seems that the sentence would then

⁴ Giere (1988, Ch. 3) argues that models are 'abstract entities', which could be also interpreted as a fiction based view of models. However, in personal communication he pointed out to me that this is not his intended view.

be about nothing, and yet we qualify such sentences as true or false. On what grounds do we do this?

These and other related concerns have led many philosophers to dismiss fictional entities. So how is appeal to something as problematic and obscure as fictional entities going to help us work through the thorny problem of scientific representation? Before turning to the details of the account that I favour (Section 4), I want to mention four reasons for believing that thinking about modelling in this way is helpful.

First, works of fiction characteristically do not portray actual states of affairs. The names of persons and objects in literary fiction characteristically do not denote real persons or objects, and there is nothing in the world of which the text of a novel is a true description.⁵ Nevertheless, fictional discourse is genuinely meaningful: readers neither make a mistake, nor are they under an illusion when they believe that they understand the contents of a novel. Yet, at the same time they are fully aware that the sentences they read when engaging with a work of fiction do not describe anything in the actual world. The same is true of modelling discourse in science. As we have seen above, scientific discourse abounds with descriptions that are meaningful yet fail to be plain descriptions of physical systems from the domain of enquiry of the scientific discipline in question.

Second, we can truly say that in David Lodge's *Changing Places* Morris Zapp is a professor of English literature at the State University of Euphoria. We can also truly say that in the novel he has a heart and a liver, but we cannot truly say that he is a ballet dancer or a violin player. Only the first of these claims is part of the explicit content of the novel, yet there is a matter of the fact about what is the case 'in the world of the story' even when claims go beyond what is explicitly stated. Whether or not claims about a story's content are correct is – somehow – determined by the text without being part of its explicit content. Such determinations are not merely decided by each reader on a whim. The situation with model-systems is the same. Model-descriptions usually only specify a handful of essential properties, but it is understood that the model-system has properties other than the ones mentioned in the description. Model-systems are interesting exactly because more is true of them than what the initial description specifies; no one would spend time studying model-systems if all there was

⁵ This is not meant to be a *definition* of fiction. A failure of reference, although typical for fiction, is neither necessary nor sufficient for a text to qualify as fiction. I come back to this point later on.

to know about them was the explicit content of the initial description. It is, for instance, true that the Newtonian model-system representing the solar system is stable and that the model-earths move in elliptical orbits; but none of this is part of the explicit content of the model-system's original specification.

Third, a fictional story not only has content that goes beyond what is explicitly stated, we also have the means to learn about this 'extra content' by using certain (usually implicit) rules of inference. It is an integral part of our response to fiction that we supplement the explicit content and fill in facts about the plot even where the text is silent. In fact, a good part of the intellectual pleasure we get from reading a novel derives from this imaginative 'filling in' of the 'missing content'. The same goes for model-systems. Finding out what is true in a model-system beyond what is explicitly specified in the relevant description is a crucial aspect of our engagement with the system. In fact the bulk of the work that is done with a model-system is usually expended on establishing whether or not certain claims about it hold true. Is the solar system stable? Do the populations of predators and prey reach some equilibrium? Do prices stabilise? These are questions we want to answer given what we know about the model and certain other rules we regard as valid in the context in which the model-system is discussed.

Fourth, sometimes we read just for pleasure, but in particular when we read serious literature we often engage in comparisons between the characters and situations in the fiction and real situations and characters with which we are familiar. We recognise aspects of the protagonist's behaviour in someone we know and suddenly begin to understand some of his behavioural patterns: we learn about the world by reading fiction. Again, this has parallels in the context of modelling, where we learn from models about the world. Once we think about models as fictions this parallel becomes salient and urges us to think about how 'knowledge transfer' from a fictional scenario to the real world takes place.

Needless to say, this list of communalities between scientific modelling and literary fiction is neither complete, nor should it be understood as suggesting that there are no important differences between the two. The purpose of this list is to make it plausible that thinking about models as alike to literary fiction is a fruitful point of departure.

In the next section I defend this conception of model-systems against its structuralist rival. Those already convinced by the fiction view can skip this section without loss and continue with Section 4 where I present a detailed formulation of the fiction view of models.

3. Strictures on Structures

Stop and rewind. Many will think that this discussion has taken a wrong turn right at the beginning and has gotten onto a path leading straight into a thicket of confusions. The wrong turn is to take talk about nonexistent systems seriously. Worse, trying to make good on this idea by working out a theory of fiction is a pilgrimage to the devil. Those whom I expect to issue such a verdict are those who hold the view that models are set theoretical structures. This view originates with Suppes (1960) and is now held by many, among them van Fraassen (1980; 1997; 2002), Da Costa and French (1990), and French and Ladyman (1997).

At the core of this approach to models lies the notion that models are structures. A structure (sometimes ‘mathematical structure’ or ‘set-theoretic structure’) S is a composite entity consisting of a non-empty set U of individuals called the domain (or universe) of the structure S and a non-empty indexed set R of relations on U . Often it is convenient to write these as an ordered triple: $S=[U, R]$.⁶

For what follows it is important to be clear on what we mean by ‘individual’ and ‘relation’ in this context. To define the domain of a structure it does not matter what the individuals are – they may be whatever. The only thing that matters from a structural point of view is that there are so and so many of them. Or to put it another way, all we need is dummies or placeholders. Relations are understood in a similarly ‘deflationary’ way. It is not important what the relation ‘in itself’ is; all that matters is between which objects it holds. For this reason, a relation is specified purely extensionally, that is, as class of ordered n -tuples and the relation is assumed to be nothing over and above this class of ordered tuples. Thus understood, relations have no properties other than those that derive from this extensional characterisation, such as

⁶ Sometimes structures are defined so that they also include operations. Although convenient in some contexts, this is unnecessary because ultimately operations reduce to relations (Boolos and Jeffrey 1989, 98-99).

transitivity, reflexivity, symmetry, etc. This leaves us with a notion of structure containing dummy-objects between which purely extensionally defined relations hold.⁷

Let us illustrate this with a simple example. Consider $S_t = [U=(a, b, c), R=(\langle a, b \rangle, \langle b, c \rangle, \langle a, c \rangle)]$, a structure consisting of a three object domain (with the objects a , b , and c) endowed with a transitive relation R , (where ' $\langle a, b \rangle$ ' is an ordered tuple expressing that R holds between a and b).⁸ In fact, the formula in the previous sentence is all we need in order to completely define the structure. It does not matter what the objects are: their materiality is immaterial. It doesn't matter whether they are books, railway bridges, or supernovae – all that is needed is that they are objects. In the same way it does not matter whether the relation R is 'greater than' or 'older than' or 'more appreciated than' – all that matters is that R holds between a and b , and b and c , and a and c , no matter what R 'in itself' is.

A view that takes model-systems in science to be structures in this sense is too austere to serve as a basis for an account of scientific modelling. Although structures do play an important role in scientific modelling, model-systems cannot be *identified* with structures. What is missing in the structuralist conception is an analysis of the 'material' character of model-systems: even perfectly spherical planets are taken to have mass, populations are taken to consist of rabbits and foxes, etc. The view of model-systems that I advocate regards model-systems as imagined physical systems, i.e. as hypothetical entities that, as a matter of fact, do not exist spatio-temporally but nevertheless have non-structural properties in the same way in which literary characters do. I will explain below in detail how to understand this claim and address the problems that it faces. The aim of this section is to argue that this is the right way of thinking about model-systems.

There are several reasons to prefer this take on model-systems over the structuralist account. The first is the evidence from scientific practice: scientists often talk about model-systems as if they were physical things. Young and Freedman, when presenting their model of the baseball in the above quote, do not say that they present a mathematical structure. Rather they

⁷ See Russell (1919, 60) for clear account of this feature of structures.

⁸ A relation is transitive iff it is true that whenever the relation holds between objects a and b , and between b and c , then it also holds between a and c . Examples for transitive relations are *more expensive than* and *taller than*; and example for a non-transitive relation is *liking* (since it may well be that a likes b , and b likes c , but a does not like c at all).

describe a hypothetical situation in which a rigid ball moves without air resistance and in the absence of other confounding factors. This way of thinking about model-systems is typical in mechanics as well as many branches of physics. And the same is true in biology. Godfrey-Smith (2006, 736-8) points out that Levins' work on population biology – as well as the models of Maynard Smith and Szathmáry's in evolutionary theory, and hence most of the work in their respective fields – is best understood as describing imagined concrete populations. Further, Godfrey-Smith adds that this way of looking at model-systems in these fields is integral to the discovery of novel phenomena and to making sense of the treatment of certain issues (e.g. the discussion of robustness in Levins), as well as to the communication of the results in books and papers, even where the models make essential use of mathematical techniques.

Closely related to this point is the fact that the fictional scenario plays a crucial role in understanding how a model relates to reality. This is best illustrated with a simple example from population dynamics.⁹ Imagine you have a newborn pair of rabbits, one male the other female, and you also have a large garden which is their habitat. You then want to know how many pairs of rabbits you will have at some later time, and so you turn to a text on population dynamics where you find a simple model (going back to Leonardo of Pisa, also known by his nickname 'Fibonacci'). The model tells you that the population at time t_n equals the population at time t_{n-1} plus the population at time t_{n-2} . According to the model, then, we have $P(t_n)=P(t_{n-1})+P(t_{n-2})$, where $P(t_n)$ is the population at time t_n and where the distance between two instants of time is the time rabbits need to mature and breed (the numbers $P(t_n)$ are known as 'Fibonacci numbers').¹⁰ Let us assume this time is one month. Thus, the model tells us that if we start with one young pair, we have five pairs after five months, eight pairs after six months, thirteen pairs after seven months, and so on.

If you are now getting excited because you figure that your rabbit population will grow really fast (after ten months you already have fifty-five pairs according to the model),

⁹ For a discussion of this example see Smith (2007, 24-29).

¹⁰ Strictly speaking this is not a structural formulation of the model, but a structural version could easily be constructed from the equation defining the Fibonacci numbers. However, since such a construction requires some setting up (as the example in Section 9 below shows) and nothing in my conclusion depends on having such a formulation, I will not dwell on this point here.

you will be disappointed. Quite soon the real number of rabbit pairs will start diverging dramatically from the value the model predicts. This may take you by surprise, but it should not if you understand the *entire* model. The above equation is not about rabbits *per se*; it is about rabbits that never die, a garden that is infinitely large and contains enough food for any number of rabbits, and rabbits that procreate at a constant rate at constant speed. This is not by any standards an accurate description of the real situation; it is a fictional scenario and $P(t_n)=P(t_{n-1})+P(t_{n-2})$ is true of this scenario. It is crucial to appreciate this fact if we want to know under what circumstances and to what extent conclusions derived in the model can be expected to bear out in the real system. Real rabbits don't live forever, but they live for some years; the garden is not infinite but large enough to provide food and shelter for about one hundred pairs; etc. So we come to the conclusion that model is probably good for about the first nine or ten months and then starts breaking down. This is important to know when using the model, but – and this is the crucial point – there is nothing in the mathematics that tells you any of this! What makes you understand the how the model relates to the world and when and where you can reasonably use it is a comparison between the fictional scenario and the real world. So the fictional scenario is an integral component of the model, and one that cannot be eliminated and replaced by structures.

Some might now reply that the fictional scenario merely plays a pragmatic role in our use of the model (whatever that means) and can therefore be eliminated in a final formulation of the model. I disagree because, as I have just outlined, the fictional scenario is essential to the functioning of the model. But irrespective of how this issue is resolved, the structuralist conception of models faces further difficulties when we think about how a model comes to be a representation of a target-system.

A structure *per se* is not about anything at all, let alone about a particular target-system; they are pieces of pure mathematics, devoid of empirical content. But a representation must possess 'semantic content' or 'aboutness'; that is, it must stand for something else. Those who take model-systems to be structures suggest connecting structures to target-systems by setting up an isomorphism between model-system and target.¹¹ Two structures $S=[U, R]$ and $S_T=[U_T, R_T]$ are isomorphic iff there exists an isomorphism between them. An isomorphism is a mapping f :

¹¹ Other suggestions include partial isomorphism, homomorphism, and embedding – nothing in what follows depends on which one of these one chooses.

$U_T \rightarrow U$ such that f is one-to-one (bijective) and it preserves the system of relations in the following sense: the elements a_1, \dots, a_n of S_T satisfy the relation R^T iff the corresponding elements $b_1=f(a_1), \dots, b_n=f(a_n)$ in S satisfy R , where R is the relation in S corresponding to R^T .

This definition of isomorphism brings a predicament to the fore: a morphism holds between two structures and not between a structure and a part of the world *per se*. In order to make sense of the notion that there is an isomorphism between a model-system and its target-system, we have to assume that the target exemplifies a particular structure. The problem is that this cannot be had without bringing non-structural features into play.

The argument for this claim proceeds in two steps (Frigg 2006, 55-56). The first is to realise that *possessing structure S* (where S is some particular structure) is a concept that does not apply unless some more concrete concepts apply as well. Hence we cannot say that a target-system has structure S unless we also say that it has certain more concrete properties as well. Let us make this more precise with the notion of one concept being more abstract than another concept.

Concept a is more abstract than concept b iff b belongs to a class B of concepts (and $a \notin B$) such that¹²

- (i) for a to apply it is necessary that at least one $b \in B$ applies, and,
- (ii) on any given occasion, the fact that $b \in B$ applies is what the applying of a on that occasion consists in.

In other words, the concepts in B are used to ‘fit out’ the abstract concept a on any given occasion. *Working*, for instance, is more abstract in this sense than *writing a letter* or *attending a meeting*. Condition (i) says that for it to be the case that I am working, I either have to write a letter, attend a meeting, or ...; if I don’t do any of these, then I am not working. Condition (ii) says that my working on a given occasion consists in, say, writing a letter. If I complain to someone that I have been writing letters all day, and he then replies ‘OK, but when did you work?’ he is either making a joke or does not get the point (namely that writing letters is working). In other words, the two conditions say that there is no such thing as working and only working.

¹² This definition is adapted from Cartwright (1999, 39).

Having structure S is like *working* in that it needs fitting out on every occasion in which it applies. It follows from the definition of a structure that for something to have structure *S* it has to be the case that *being an object* must apply to some of its parts, and *standing in a relation R* (where *R* is one of the relations of *S*) must apply to these. These concepts are abstract relative to more concrete concepts. Let us take relations first. Recall that relations are defined purely extensionally and hence have nothing but logico-mathematical properties such as transitivity. Consider, then, *standing in a transitive relation*. 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 *standing in a transitive relation* applies to two objects only if either *greater than, or older than, or ...* applies to them as well. We cannot have the former without the latter: 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. So *standing in a transitive relation* is abstract relative to more concrete concepts like *being hotter than* and, hence there simply is no such thing in the physical world as a relation that is nothing but transitive.

Similarly for objects. What is needed for something to be an object is not an easy question, and an answer depends on the relevant context as well as the kinds of things we are dealing with (medium size physical objects like tables, social entities such as families, etc.). But nothing in the world is such that the only property it possesses is ‘objectness’; whatever the circumstances, some other concepts must apply to it for it to be the case that it is an object. For instance, a medium size physical object has an identifiable shape which sets it off from the environment, which implies that it is coloured, has a certain texture, etc. If none of this was the case, we just would not have a medium size physical object.

The crucial point in all this is that the more concrete concepts that are needed to ground structural claims are not structural themselves. *Being a transitive relation* is structural, *being taller than* is not, as becomes clear from what has been said about structures above. In other words, structural claims ride on the back of non-structural claims.

This by itself would not have to worry the structuralist who claims that model-systems are structures. He could point out that although, as the above argument shows, structures are grounded in something else (which is non-structural), it is the structural features of reality that models relate to and that therefore models are structures. The problem with this response – and this is the second step of the argument – becomes apparent when we realise that the descriptions we choose to fit out abstract structural claims almost never are true descriptions of the target systems. The above examples make this sufficiently clear. The structure on which the formal treatment of the solar system is based is not fitted out by a realistic description of the solar system, but by a description that takes planets to be ideal spheres with homogenous mass distributions gravitationally interacting only with each other and nothing else. Similarly, the structure on which the calculations of the population sizes is based does not attach to a realistic description of animal life and so on. So the structural claims that give rise to the equations that we study when dealing with a problem at hand (at least in the overwhelming majority of cases) are not true descriptions of the target system, and hence the target does not have the structure at stake.¹³

Hence, taken literally, descriptions that ground structural claims (almost always) fail to be descriptions of the intended target system. Instead, they describe a hypothetical system which is distinct from the target system. This has unfortunate consequences for the structuralist. If the descriptions employed to attribute a structure to a target system were just plain descriptions of that system, then the claim that model-systems are just structures would appear at least *prima facie* plausible. But once we acknowledge that these descriptions describe hypothetical systems rather than real target systems, we also have to acknowledge that hypothetical systems are an important part of the theoretical apparatus we employ, and that they therefore have to be included in our analysis of how scientific modelling works. This can, of course, be done in different ways. My suggestion is that these hypothetical systems in fact *are* the model-systems. I therefore I reserve the term ‘model-system’ for the hypothetical physical entities described by the descriptions we use to ground structural claims; I refer to

¹³ This is what Downes has in mind when he says that there is no empirical system corresponding to the equation of the ideal pendulum (1992, 145), and what Thomson-Jones (2007) emphasises when he points out that science is full of ‘descriptions of missing systems’; in a different ways the same point is also made by Cartwright (1983, Ch. 7) who emphasises that we have to come up with a ‘prepared description’ of the system in order to make it amenable to mathematical treatment.

the relevant structures as ‘model structures’. This facilitates the analysis in what follows, but ultimately nothing hangs on this choice; one could just as well say that model-systems are composite entities consisting of a hypothetical and a structural system. What does matter, however, is that we acknowledge that scientific modelling indeed involves such hypothetical systems.¹⁴

At least some proponents of structuralist conception will reject this argument.¹⁵ The bone of contention is what model-systems represent. So far I have assumed that a model-system represents a piece of the real world, for instance the solar system or a population of rabbits. This, so the objection goes, is the wrong point of departure since models don’t represent systems in this sense. What a model-system ultimately represents is a *data model*, not an object of some sort. Data are what gather in experiments. When observing the motion of the moon, we take choose a coordinate system and observe the position of the moon in this coordinate system at consecutive instants of time. We then write down these observations. This can be done in different ways. We can simply write a list with the coordinates of the moon at certain instants of time; we can draw a graph consisting of various points standing for the position of the moon at different times; or we can choose yet another form of taking down the data. The data thus gathered are called the *raw* data. The raw data then undergo a process of cleansing, rectification and regimentation: we throw away data points that are obviously faulty, take into consideration what the measurement errors are, take averages, etc. Often (but not always) the aim of this process is to fit a smooth curve through the various data points so that the curve satisfies certain theoretical desiderata (having minimal least-square-distance from the actual data points). The end result of this process is a so-called *data model*.

¹⁴ One could try to avoid the commitment to hypothetical systems by renouncing a literal understanding of the relevant descriptions and arguing that it does not follow from the fact that descriptions are poor or highly idealised that they are not descriptions of the target at all; it just means that they are idealised descriptions. This move is of no avail. Being an idealised description is not a primitive concept and it calls for analysis. On the most plausible analysis, *D* is an approximate description of object *O* iff what *D* literally describes is in some relevant sense an idealisation of *O*. But what *D* literally describes is a hypothetical system, and so we find ourselves back where we started.

¹⁵ The German structuralists explicitly acknowledge the need for a concrete description of the target-system (Balzer, Moulines, and Sneed 1987, 37-38). Moreover, they consider these ‘informal descriptions’ to be ‘internal’ to the theory. Unfortunately they do not say more about this issue. Nevertheless, it is important to emphasise that there is no conflict between structuralism thus construed and the view developed in this chapter; in fact they can be seen as complementary.

The claim then is that model-systems do not represent parts of the world (like the earth and the sun), but rather data-models that have been constructed from observations made on these parts of the world. So what a model of the motion of planet earth is about is not the earth itself, but the smooth curve that we have fitted through the data gained when observing the motion of the earth. In this vein van Fraassen declares that ‘[...] 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.’ (2002, 164).¹⁶ In brief, the suggestion is that representation be explicated in terms of setting up an isomorphism between the model-system (on this view a structure) and the data model. This move indeed renders the above argument obsolete since data models are mathematical entities and as such can be considered to have a well-defined structure.¹⁷

This suggestion is wrong because it is descriptively inadequate: it is not the case that models represent data. This point is not new. It has been argued by Bogen and Woodward (1988) and Woodward (1989), and has recently been reiterated in different guise by Teller (2001).¹⁸ In essence I agree with these authors; however, my focus differs slightly from theirs and I present the subject matter in a way that suits my needs.

In *nuce*, Bogen’s and Woodward’s point is that science is not about data; it is about phenomena. A theory about the melting point of lead is not about the data we gather when we find out at what temperature lead melts; it is about the melting of lead itself. This carries over to models: models do not represent data. In fact, 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.

Let me illustrate this with an example from Bogen and Woodward: the discovery of weak neutral currents (*ibid.*, 315-18). What the model at stake consists of is particles: neutrinos,

¹⁶ See also van Fraassen (1980, 64; 1989, 229; 1997, 524) and French (French 1999, 191-192).

¹⁷ There is an exegetic question here. Although structuralists certainly suggest that representation *is* data matching, they never explicitly say so. I here explore the stronger version of the view on which representation indeed *consists in* data matching since the weaker version, on which data matching is distinct from representation, does not provide a viable criticism of the above argument from abstractness.

¹⁸ McAllister (1997) presents an antirealist critique of Bogen and Woodward. But his concern is orthogonal to mine: even if one construes phenomena in an antirealist way they turn out to be more than just data.

nucleons, the Z^0 , and so on, along with the reactions that take place between them.¹⁹ Nothing of that, however, shows in the relevant data. What was produced at CERN in Geneva were 290000 bubble chamber photographs of which roughly one hundred were 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 gap between observable outcomes of experiments and anything one might call a substructure of a model of neutral currents.²⁰

¹⁹ 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.

²⁰ To underwrite this claim consider the following example. Parallel to the research at CERN, the NAL in Chicago 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 the data gathered had nothing in common, they were meant to provide evidence for the same theoretical model. But the model 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. The model is not idiosyncratic to a special experimental context in the way the data are, and therefore it is not surprising that the model does not contain a substructure that could plausibly be claimed to be isomorphic to the data. The model represents an entity – weak neutral currents – and not data used in its discovery.

But 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, and for the fact that the model is a (more or less) faithful representation of what is happening in the world. 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.

In sum, understanding the fictional scenario of which the formal apparatus of a model is literally true is essential to understanding and using a model. Furthermore, one has to recognise that structures cannot be connected to anything in the world without the mediation of non-structural concepts, and attempts to bypass this conclusion by appeal to data models fails.

4. Model-Systems and Imagination

So far, I have argued that model-systems are best understood as akin to characters and objects of literary fiction. However, as I have indicated above, fictional entities are beset with philosophical problems (see Friend (2007) for a discussion of these) and hence explaining models in terms of fiction hardly seems to be progress. Hence the burden of proof is on the side of the proponent of the fiction view, who has to show that there is a workable conception of fiction that serves the needs of a theory of scientific modelling. Developing such a view is the aim of this section.²¹ This involves a lengthy discussion of philosophical subtleties that at first may seem peripheral to the concerns of scientific modelling. I appeal to the forbearance of the reader and promise that this effort is not in vain. For one, without a tenable conception of fiction, the fictions view is without foundation, and the only way to prove that it stands firm is to explicitly formulate a tenable account of fiction. For another, one of the results of this excursion into the philosophical jungles of fiction is the distinction it allows us to draw between two different conceptions of representation, p-representation and t-representation. This distinction, I think, is crucial to understanding how scientific modelling works, and a failure to keep the two separate has led to considerable confusion.

²¹ This section and the next are based on my (2009).

What do we expect from an account of fiction in order for it to be able to serve as the foundation of the fiction view of model-systems? I think it has to provide responses to five questions (Q1-Q5) and to satisfy two meta-theoretical criteria (C1-C2). These questions and criteria are as follows:

(Q1) *Identity conditions*. When are two model-systems identical? This question is pressing because unlike in the context of literature, where we can point to canonical texts and authors' intentions, model-systems in science are often presented by different authors (in different papers or textbooks) in different ways. Nevertheless, many different descriptions are actually meant to describe the same model-system. Under what circumstances is that the case? That is, when are the model-systems specified by different descriptions identical?

(Q2) *Attribution of properties*. In the previous section I have argued that model-systems have 'physical', 'concrete', or 'material' properties. As the scare-quotes indicate, there is something problematic about this claim. In fact, it has even been claimed that such statements are outright contradictory because abstract objects like the ideal pendulum cannot have the same properties as concrete physical systems (Hughes 1997, 330). How is it possible for a model-system to have 'material' properties if model-systems do not exist in space and time? What sense can we make of statements like 'the ball is charged' or 'the population is isolated from its environment' if there are no balls and populations?

(Q3) *Comparative statements*. As we have seen above, comparing a model and its target-system is essential to many aspects of modelling, and it plays a crucial role in the account of representation developed below. We customarily say things like 'real agents do not behave like the agents in the model' and 'the surface of the real sun is unlike the surface of the model sun'. How can we compare something that does not exist with something that does? Likewise, how are we to analyse statements that compare features of two model-systems with each other like 'the agents in the first model are more rational than the agents in the second model'?

(Q4) *Truth in model-systems*. There is right and wrong in a discourse about model-systems. It is true that the population in Fibonacci's model never decreases and it is wrong that the earth in Newton's model moves on parabolic orbit. But on what basis are claims about a model-system qualified as true or false, in particular if the claims concern issues about which the

description of the system remains silent? What we need is an account of truth in model-systems, which, first, explains what it means for a claim about a model-system to be true or false and which, second, draws the line between true and false statements at the right place (for instance, an account on which all statements about a model-systems come out false would be unacceptable).

(Q5) *Epistemology*. We do investigate model-systems and find out about them; truths about the model-system are not forever concealed from us. In fact, we engage with model-systems because we want to explore their properties. How do we do this? How do we find out about these truths and how do we justify our claims?

(C1) *Naturalism*. The account we offer in response to the above issues should be able to make sense of scientific practice. That is, it should be able to explain how scientists build models and how they reason about them.

(C2) *Metaphysical commitments*. The metaphysics of fictional entities is an issue fraught with controversy. For this reason we need to know what kind of commitments we incur when we understand model-systems along the lines of fiction, and how these commitments, if any, can be justified. However, it is not, in my view, a condition of adequacy that the account we propose be metaphysically parsimonious. As a matter of fact, the account I develop below eschews commitment to fictional entities, but this is accidental, as it were. To say it a different way, it just so happens that the theory that provides the most convincing answers to the above questions is also metaphysically parsimonious; but if it had turned out that a metaphysically substantial theory (i.e. one that is committed to fictional entities) had provided the best answers, then we should have chosen that theory. In other words, I think that accounts of fictional entities should not be dismissed merely on the grounds of being metaphysically ‘thick’. That I dismiss such accounts has to do only with their failure to answer other questions in a satisfactory way.²²

²² For want of space I cannot discuss competing approaches. In a nutshell, their problems seem to be the following. The paraphrase account (Russell 1905) does not offer a workable theory of truth in fiction (Crittenden 1991, Ch. 1). The neo- Meinongean view (Parsons 1980) runs into difficulties with incompleteness (Howell 1979, Sec. 1) and as a consequence does not offer a satisfactory answer to (I5). Finally, Lewis’ (1978) account is too permissive about what counts as true in a fictional context (Currie 1990, Sec. 2.3; Lamarque and Olsen 1994, Ch. 4).

That said, it is the contention of this chapter that Kendall Walton's (1990) pretence theory of fiction best fits this bill.²³ The next section provides a brief introduction to this theory, and Section 5 outlines the responses that we get from this theory to (Q1) – (Q5) and (C1)-(C2). The point of departure of this view is the capacity of humans to imagine things.²⁴ Sometimes we imagine something without a particular reason. But there are cases in which our imagining something is prompted by the presence of a particular object, in which case this object is referred to as a 'prop'. 'Object' has to be understood in the widest sense possible; anything capable of affecting our senses can serve as a prop. An object becomes a prop due to the imposition of a rule or 'principle of generation' (p. 38), prescribing what is to be imagined as a function of the presence of the object. If someone imagines something because he is encouraged to do so by the presence of a prop he is engaged in a game of make-believe. Someone who is involved in a game of make-believe is pretending; so 'pretence' is just a shorthand way of describing participation in such a game (p. 391) and has (in this context) nothing to do with deception (p. 392). The simplest examples of games of make-believe are cases of child's play (p. 11). In one such case, stumps may be regarded as bears and a rope put around the stump may mean that the bear has been lassoed; or pointing the index finger at someone and saying 'bang' may mean that the person has been shot.

A prop becomes a prompter if some notices the prop and as a result starts engaging in a rule-guided imaginative activity. The set of prompters and the set of props overlap, but neither is a subset of the other. For one, a prop that is never perceived by anybody and hence never causes anybody to imagine something is not a prompter (but still a prop). For another, an object can prompt imaginations without being part of a game of make-believe (i.e. in the absence of rules of generation), for instance when we see faces in the clouds and imagine how these faces talk to each other. Even within a game we can make errors (e.g. mistakenly take a mole heap for a stump and then say that it is a bear), in which case the mole heap is a prompter (because it prompts imaginings) but it is not a prop (because there is not a rule).

²³ Strictly speaking, Walton (1990) restricts the use of 'pretence' to verbal (or more generally behavioural) participation, which does not include the activity of someone reading on his own. However, it has become customary to use 'pretence' as synonymous with 'make-believe' and I stick to this wider use in what follows.

²⁴ I here discuss pretence theory as it is presented by Walton (1990); Currie (1990) and Evans (1982, Ch. 10) develop different versions. Parenthetical references in the text of this and the following section are to Walton's book.

Pretence theory considers a vast variety of different props ranging from novels to movies, from paintings to plays, and from music to children's games. In the present context I only discuss the case of literature. Works of literary fiction are, on the current account, regarded as props as they prompt the reader to imagine certain things. By doing so a fiction generates its own game of make-believe. This game can be played by a single player when reading the work, or by a group when someone tells the story to the others.

Some rules of generation are *ad hoc*, for instance when a group of children spontaneously imposes the rule that stumps are bears and play the game 'catch the bear'. Other rules are publicly agreed on and hence (at least relatively) stable. Games based on public rules are 'authorized'; games involving *ad hoc* rules are 'unauthorized'.

By definition, a prop is a representation if it is a prop in an *authorised* game. On this view, then, stumps are not representations of bears because the rule to regard stumps as bears is an *ad hoc* rule that is neither shared by others in the society nor stable over time (stumps may not be props to other people and even the children playing the game now may regard them as elephants on the next walk). However, *Hamlet* is a representation because everybody who understands English is invited to imagine its content, and this has been so since the work came into existence. Within pretence theory 'representation' is used as a technical term. Representations are not, as is customary, explained in terms of their relation (e.g. resemblance or denotation) to something beyond themselves; representations are things that possess the social function of serving as props in authorised games of make-believe (I will come back to this point below).

Props generate fictional truths by virtue of their features and principles of generation. Fictional truths can be generated directly or indirectly; directly generated truths are 'primary' and indirectly generated truths are 'implied' (p. 140). Derivatively, one can call the principles of generation responsible for the generation of primary truths 'principles of direct generation' and those responsible for implied truths 'principles of indirect generation'. The leading idea is that primary truths follow immediately from the prop, while implied ones result from the application of some rules of inference. When little Jimmy sees a stump and shouts 'here is a bear' this is a direct truth because it follows from fact that there is a stump and the direct rule 'stumps are bears', which is constitutive of the game. The boys may then stay away from the bear because they think the bear is dangerous and might hurt them. This fictional truth is

inferred because it does not follow from the basic laws of the game that stumps are bears, but from the additional principle that bears in the game have the same properties as real bears.

The distinction between primary and inferred truths is also operative in literary fiction. The reader of *Changing Places* reads that Zapp ‘embarked [...] on an ambitious critical project: a series of commentaries on Jane Austen which would work through the whole canon, one novel at a time, saying absolutely everything that could possibly be said about them.’ The reader is thereby invited to imagine the direct truth that Morris Zapp is working on such a project. She is also invited to imagine that Zapp is overconfident, arrogant in an amusing way, and pursues a project that is impossible to complete. None of this is explicitly stated in the novel. These are inferred truths, which the reader deduces from common knowledge about academic projects and the psyche of people pursuing them.²⁵ What rules can legitimately be used to reach conclusions of this sort is a difficult issue fraught with controversy. I will return briefly to it below; for the time being all that matters is that there are such rules, no matter what they are.

This framework, has the resources to explain the nature of model-systems. Typically, model-systems are presented to us by way of descriptions, and these descriptions should be understood as props in games of make-believe., These descriptions usually begin with expressions like ‘consider’ or ‘assume’ and thereby make it clear that they are not descriptions of fact, but an invitation to ponder – in the present idiom, imagine – a particular situation. Although it is often understood that this situation is such that it does not occur anywhere in reality, this is not a prerequisite; models, like literary fictions, are not *defined* in contrast to truth. In elementary particle physics, for instance, a scenario is often proposed simply as a suggestion worth considering. Only later, when all the details are worked out, the question is asked whether this scenario bears an interesting relation to what happens in nature, and if so what the relation is.²⁶

²⁵ The distinction between primary and inferred truths is not always easy to draw, in particular when dealing with complex literary fiction. Walton also guards against simply associating primary truth with what is explicitly stated in the text and inferred ones with what follows from them (see Walton (1990, Ch. 4) for a discussion). For the purpose of the present discussion these subtleties are inconsequential.

²⁶ For an accessible account of particle physics that makes this aspect explicit see Smolin (2007), in particular Ch. 5

The ‘working out’ of the details usually consists in deriving conclusions from the primary assumptions of the model and some general principles or laws that are taken for granted. For instance, we derive that the earth moves in an elliptical orbit from the basic assumptions of the Newtonian model and the laws of classical mechanics. This is explained naturally in the idiom of pretence theory. What is explicitly stated in a model description (that the model-earth is spherical, etc.) are the primary truths of the model, and what follows from them via laws or general principles are the implied truths; the principles of direct generation are the linguistic conventions that allow us to understand the relevant description, and the principles of indirect generation are the laws that are used to derive further results from the primary truths.

We can now address the above questions. The attribution of certain concrete properties to models (Q2) is explained as it being fictional that the model-system possesses these properties. To say that the model-population is isolated from its environment is just like saying that Zapp drives a convertible. Both claims follow from a prop together with rules of generation. In other words, saying that a hypothetical entity possesses certain properties involves nothing over and above saying that within a certain game of make-believe we are entitled to imagine the entity as having these properties. For this reason there is nothing mysterious about ascribing concrete properties to nonexistent things, nor is it a category mistake to do so.

Let us now discuss the issue of truth in model-systems (Q4), which will also provide us with solutions to the other open questions. The question is: what exactly do we assert when we qualify ‘Zapp drives a convertible’ as true in the fiction while ‘Zapp drives a Mini Cooper’ as false?²⁷ To begin with, it is crucial to realise that there are three different kinds of statement in connection with fiction, and that these require a different treatment when it comes to the questions of truth; I refer to these as intrafictional, metafictional, and transfictional statements.²⁸ For someone sitting in an armchair reading *Changing Places* ‘Morris jumped into the paternoster on the downside’ is an intrafictional statement because the reader is

²⁷ There is controversy over this issue even within pretence theory. It is beyond the scope of this paper to discuss the different proposals and compare them to one another. In what follows I develop an account of truth in fiction that is based on elements from different theories and that is tailored towards the needs of a theory of model-systems.

²⁸ All theories of fiction acknowledge this distinction. My terminology is adapted from Currie (1990, Ch. 4) who speaks about the ‘fictive’, ‘metafictive’ and ‘transfictive’ use of fictional names.

involved in playing the game defined by the novel and imagines that the sentence's content is the case. Someone who read the novel a while ago and asserts in discussion with a friend that Zapp jumped into a paternoster makes a metafictional statement because he is talking about the fiction. If he then also asserts that Zapp, his quirks notwithstanding, is more likeable than any literature teacher he ever had or that Zapp is smarter than Candide, he makes transfictional statements as he is comparing Zapp to a real person and a character in another fiction.²⁹

Intrafictional propositions are made within the fiction and we are not meant to believe them, nor are we meant to take them as reports of fact; we are meant to imagine them. Although some statements are true in the fiction as well as true *tout court* ('1968 was the year of student revolts' is true and true in *Changing Places*), we often qualify false statements as true in the fiction ('Zapp is a literary theorist' is false because there is no Zapp) and true statements as false in the fiction ('white light is composed of light of other colours' is false in Goethe's *Faust*). So truth and truth in fiction are distinct; in fact, truth in fiction is not a species of truth at all (p. 41). For this reason it has become customary when talking about what is the case in a fiction to replace locutions like 'true in the fiction' or 'true in a fictional world' by the term of art 'being fictional'; henceforth ' $F_w(p)$ ' is used as an abbreviation for 'it is fictional in work w that p ', where p is a placeholder for an intrafictional proposition like 'Zapp pursues an impossible project'.³⁰

The question now becomes: when is p fictional in w ? Let the w -game of make-believe be the game of make-believe based on work w , and similarly for ' w -prop' and ' w -principles of generation'. Then, p is fictional in w iff p is to be imagined in the w -game of make-believe (p. 39). In more detail:

²⁹ Notice that while transfictional statements are recognisable by the presence of terms that are foreign to the work under discussion, intrafictional and metafictional statements are recognisable as such only as a function of the context in which they appear. There are also statements that are difficult to classify. As these typically involve emotional reactions on the part of the reader to the novel (halfway through the book a reader exclaims 'I fear the worst for Zapp'), they need not occupy us here.

³⁰ I here follow Currie (1990, Ch. 2) and assume that sentences like 'Zapp drives a convertible' express propositions, something that Walton denies (p. 391). This assumption greatly simplifies the statement of truth conditions for fictional statements, but nothing in the present paper hangs on it. Essentially the same results can be reached only using sentences and pretence (see pp. 400-405).

p is fictional in work w iff the w -prop together with the w -principles of generation prescribes p to be imagined

This analysis alleviates worries about the (alleged) subjectivity of imaginings. In common parlance, ‘imagination’ has subjective overtones, which might suggest that an understanding of models as imagined entities makes them subjective because every person imagines something different. This is not so. In pretence theory, imaginations in an authorised game of make-believe are sanctioned by the prop itself and the rules of generation, both of which are public and shared by the relevant community. Therefore, someone’s imaginings are governed by intersubjective rules, which guarantee that, as long as the rules are respected, everybody involved in the game has the same imaginings. So, not only do all participants in the game *de facto* imagine the same things (which could also be the result of happenstance), but they do so because they participate in a rule-governed activity. What is more, participants *know* that they do; they know that they are participants in an authorised game and as long as they trust that the others play by the rules they can trust that other have the same imaginings.

Furthermore, for a proposition to be fictional in work w it is not necessary that it is actually imagined by anyone: fictional propositions are ones for which there is a prescription to the effect that they *have to be imagined* (p. 39), and whether a proposition is to be imagined is determined by the prop and the rules of generation. Hence, props, via the rules of generation, make propositions fictional independently of people’s actual imaginings (p. 38), and for this reason there can be fictional truths that no one knows of. If there is a stump hidden behind a bush, unknown to those playing the game, it is still fictional that there is a bear behind the bush; the prop itself and the rules of generation are sufficient to generate this fictional truth.

With this in place we can now also render concept of a ‘fictional world’ or ‘world of a fiction’ precise: the world of work w is the set of all propositions that are fictional in w .³¹

This analysis of truth in fiction carries over to model-systems one to one simply by replacing p by a claim about the model, w by the description of the model-system, and w -principles of generation by the laws and principles assumed be at work in the model. For instance, ‘the

³¹ Fictional worlds thus defined are rather different from possible worlds as used in modal logic, the most significant difference being that the former are incomplete while the latter are not. See Currie (1990, 53-70) for a discussion of possible worlds and fiction.

solar system is stable' is true in the Newtonian model of the solar system systems iff the description of the system together with the laws and principles assumed to hold in the system (the laws of classical mechanics, the law of gravity, and some general assumptions about physical objects) imply that this is the case. This gives us a straightforward answer to the question about identity conditions (Q1): two models are identical iff the worlds of the two models – the set of all propositions that are fictional in the two models – are identical.³²

Metafictional propositions make genuine claims that can be true or false in the same way in which claims about chairs and tables can be true or false. But how can such statements be true if the singular terms that occur in them have no referents? A solution emerges when we realise that statements like 'Zapp is a professor' are ellipses for 'in *Changing Places*, Zapp is a professor'. So when we metafictionally assert p , what we really assert is 'in work w , p ' (p. 397). Asserting that something is the case in a work of fiction is tantamount to asserting that it is fictional in that work. Hence asserting 'in work w , p ' amounts to asserting ' p is fictional in work w ', which in turn is equivalent to 'it is fictional in work w that p '. The last sentence is, of course, just $F_w(p)$. Hence metafictionally asserting p amounts to asserting $F_w(p)$. The truth condition for this assertion follows from what has been said above:

$F_w(p)$ is true iff p is fictional in w , which in turn is the case iff the w -prop and together with the w -principles of generation prescribes p to be imagined.

Derivatively, p , when uttered as a metafictional claim, is true iff p is fictional when uttered as an intrafictional claim.³³ In sum, once we understand that a metafictional claim has to be prefixed by 'In fiction w ', and hence has the structure $F_w(p)$, the truth of the claim is determined by appeal to the w -game of make-believe. Again, this analysis translates to scientific statements without further ado.

Transfictional propositions pose a particular problem because they – apparently – involve comparisons with a nonexistent objects, which does not seem to make sense: we cannot

³² An interesting consequence of this identity condition is that not all models with the same prop are identical, because they can operate with different rules of indirect generation. This is the case, for instance, when the 'same model' is treated first classically and then quantum mechanically; on the current view, the classical and the quantum model are not identical.

³³ In some places Walton ties the truth of such statements to *authorised* games (e.g., p. 397-8). This restriction seems unnecessary as the analysis works just as well for unauthorized games.

compare someone with Zapp if there is no Zapp. Different authors have offered very different solutions to this problem.³⁴ Fortunately we need not deal with the problem of transfictional statements in its full generality because the transfictional statements that are relevant in connection with model-systems are of a particular kind: they compare features of the model-systems with features of the target-system. For this reason, transfictional statements about model-systems should be read as prefixed with a clause stating what the relevant respects of the comparison are. This allows us to rephrase comparative sentences as comparisons between properties rather than objects, which makes the original puzzle go away.

Crucially, then, truth conditions for transfictional statements in the context of scientific modelling come down to truth conditions for comparative statements between properties, which are unproblematic in the current context (for the problems that attach to them have nothing to do with issues surrounding fictional discourse). For instance, when I say ‘my friend James is just like Zapp’ I am not comparing my friend to a nonexistent person. What I am asserting is that both James and Zapp possess certain relevant properties (Zapp possesses properties in the sense explained above) and that these properties are similar in relevant ways. Likewise, when I say that the population of rabbits in a certain ecosystem behaves very much like the population in the Fibonacci model, what I assert is that these populations possess certain relevant properties which are similar in relevant respects. What these relevant properties are and what counts as being similar in relevant respects may well depend on the context. But this is not a problem. All that matters from a semantic point of view is that the apparent comparison with a nonexistent object eventually comes down to the unproblematic comparison of properties. Further, the statement making this comparison is true iff the statement comparing the properties with each other is true. Obviously, statements comparing two nonexistent objects are analysed in exactly the same way.

These insights provide us with answers to (Q3) and (Q4). And what is more, this take on truth also provides us with an answer to the question about the epistemology of models (Q5): we investigate a model by finding out what follows from the primary truths of the model and the rules of indirect generation. This seems to be both plausible and in line with scientific practice

³⁴ Lamarque and Olsen (1994, Ch. 4), for instance, solve the problem by introducing characters. Walton, by contrast, renounces the commitment to characters and instead analyses transfictional statements in terms of unauthorized games (pp. 405-416).

because a good deal of the work that scientists do with models can accurately be described as studying consequences of the basic assumptions of the model – so can tick off (C1) as well.

What metaphysical commitments do we incur by understanding models in this way? The answer is: none. Walton's theory is antirealist in that it renounces the postulation of fictional or abstract entities, and hence a theory of scientific modelling based on this account is also free of ontological commitments. This, of course, is not a refutation of metaphysically less parsimonious views such as Meinong's, and there may be reasons to eventually prefer such a view over an antirealist one. The point to emphasise here is that whatever these reasons may be, the needs of science are not one among them.

This concludes the discussion of the conditions of adequacy of the last section, and I hope to have made it plausible that the framework of pretence theory provides convincing responses to the issues that arise in connection with model-systems.

With this in place, we can now distinguish two different kinds of representation, which will be important in understanding scientific representation. As mentioned above, pretence theory *defines* a representation to be a prop in an authorised game of make-believe. On this view, the text of a novel and the description of a model-system are representations. Derivatively one can then say that props represent the imaginings they prescribe. Although this is a common use of 'representation', the term is used rather differently in both science and philosophy of science where it is taken to denote a relation between the model-system and its target (and, depending on one's views about representation, also other relata such as users and their intentions). But far from being in conflict with each other, these two notions of representation are actually complementary – I will turn to this point in the next section. For now it is just important not to get them mixed up, and for this reason I call the former 'p-representation' ('p' for 'prop') and the latter 't-representation' ('t' for target).³⁵ Using this idiom, pretence theory (as presented in this section) can be understood as an analysis of p-representation. This leaves pending an analysis of t-representation, to which I turn in Section 6 below. I defer this

³⁵ A more intuitive choice of terminology would be to refer the term 'representation' for what I here call t-representation, and refer to p-representation as 'presentation'. However, since this would stand in conflict with the use of 'representation' in pretence theory I stick to the somewhat less elegant terminology of p- and t-representation.

task because I first want to summarise where we stand and formulate a consistent overall picture of scientific modelling, which is the aim of Section 5.

5. The Anatomy of Scientific Modelling

We have analysed model-systems in terms of imagined objects and distinguished two different representational relations, p-representation (which holds between a prop and the imaginings that it mandates) and t-representation (which holds between a thus imagined system and a target-system in the world). Using these notions, the two acts mentioned in the introduction can be described as, first, introducing a p-representation specifying an imagined object and, second, claiming that this imagined object t-represents the relevant target-system.

Putting all this together we obtain a general picture of scientific modelling. This picture is schematically illustrated in Figure 1.

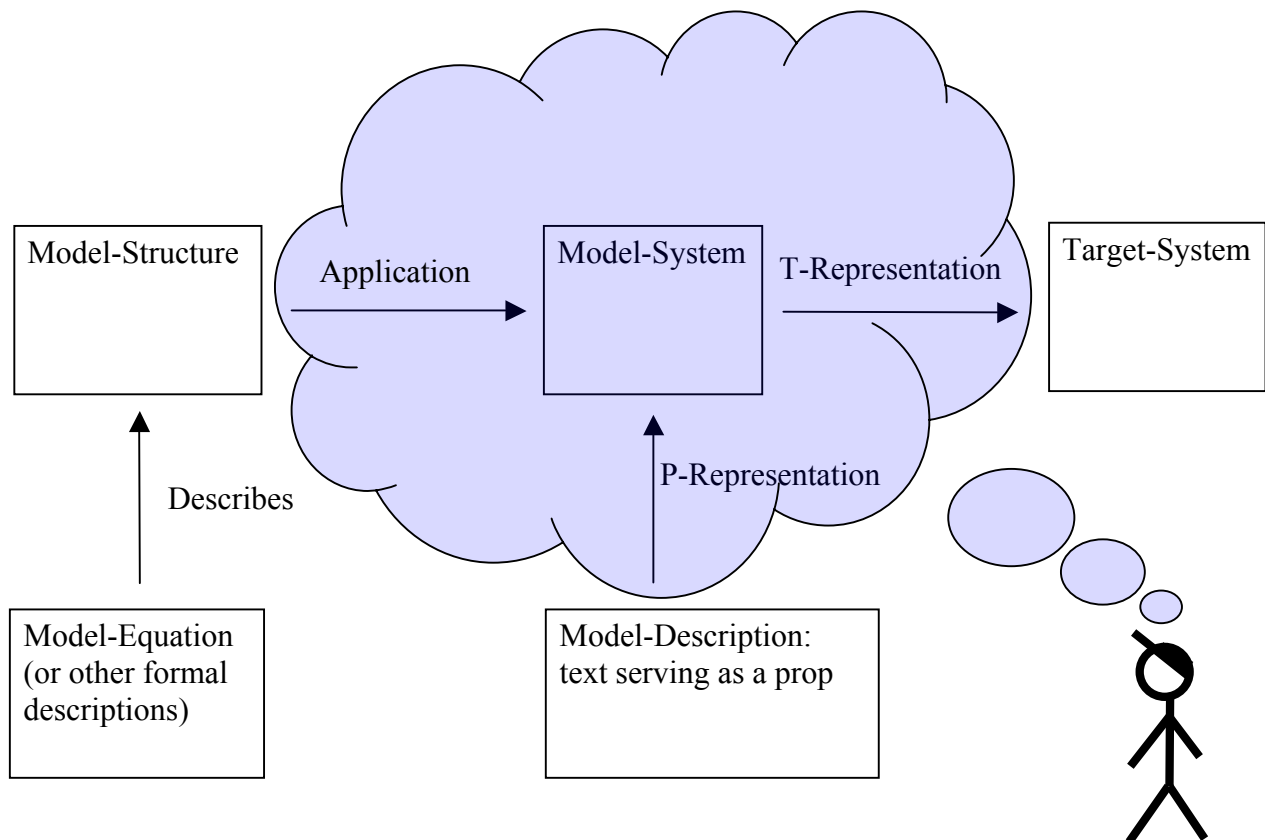


Figure 1 – The elements of scientific modelling.

The boxes in the middle and on the right emerge from the above discussion and don't need further explanation. Not so the boxes on the left. These account for the use of mathematics. How mathematics applies to something non-mathematical is a time-honoured philosophical puzzle, and much has been written about it. However, since this is somewhat peripheral to the concerns of this chapter, I will not discuss this issue further and merely put the relevant boxes into the diagram for the sake of completeness. A discussion of the issue of the applicability of mathematics can be found in Shapiro (2000).

Let me then add some points about this diagram by way of clarification and explanation. First, there is a temptation to respond to this suggestion by saying: 'yes, fine, but where in this scheme is *the model*?' There is no single answer to this question. With the exception of the target-system itself, every part of the above schema (and every combination of parts!) legitimately may be, and sometimes is, referred to as 'model', which is why I tried to avoid the term altogether. Once it is acknowledged that scientific modelling involves all the above elements, the determination of which one of these we call 'the model' is inconsequential. As long as one is aware of this we can choose terminology as we please.

Second, this picture of scientific modelling is independent of how one understands the relation between models and theories. The model-structure in this diagram is assumed to be a structure used in the treatment of a particular concrete system, and not a general structure. It is, for instance, the structure of the harmonic oscillator, the two-body system, or a conical spinning top on a frictionless plane; it is not Newtonian Mechanics, Quantum Mechanics, Fluid Mechanics, or General Relativity (in a structural rendition). This leaves open the question of how these specific structures relate to overarching theories. In particular, this picture is compatible with the semantic view of theories that would take the model-structure to belong to a family of structures which forms a theory (van Fraassen 1980). This view agrees also with the German structuralist picture that construes model-structures as being the result of a process of specification and restriction of a general theory (Balzer, Moulines, and Sneed 1987), and a view that denies that there is any straightforward connect between models and theories (Morgan and Morrison 1999).

Third, this diagram has no temporal connotations and there is no view implicit in it about what comes first in the process of the construction of a model. Sometimes we start with a

fictional scenario; sometimes we start with an equation we think might be useful; sometimes we have a clear strategy for t-representation in mind right from the start, and sometimes we just ‘try out’ something and worry later about how the model relates to the world. It is not even assumed that all parts of the diagram are belaboured by the same scientist. In particular when it comes to large and complex models (such as climate models), different groups may take care of different parts of the model (e.g. one group may develop mathematical tools and another one takes care of their application to the concrete problem at hand). In brief, this picture of modelling is compatible with any view one wants to take on the actual process of model construction and the division of labour therein.

Fourth, there is a time-honoured problem about how it is possible that we can represent something that does not exist. How can we represent Santa Claus if there is no Santa Claus? More pertinently, how can we have models representing in great detail mechanical properties of the ether if there is no ether? Thinking about modelling in the way I have proposed makes this problem go away at once, since it becomes clear that equivocating on ‘representation’ is the root of the puzzle. On the one hand, we take representation to be a relation between a picture or model and an item in the real world (which does not exist). On the other hand, it assumes representation to be the infliction of mental content in an observer when she looks at a picture or reads the description of a model (which is, of course, real). This is exactly the distinction between t-representation and p-representation. Santa Clause pictures and ether models do not t-represent because there is no Santa Claus and no ether. But Santa Claus pictures p-represent in that they become props in a game of make believe leading us to imagine all kind of things about a bearded old man in a red outfit bringing gifts, and a description (or graphical representation) of an ether model leads us to imagine a fictional model-system. Once we recognise the distinction between p-representation and t-representation, the problem evaporates.³⁶

³⁶ Model-systems without targets (and hence without t-representation) not only play a role when explaining failures; they are also important as means to explore certain technical tools, in which case they are often referred to as ‘probing models’, ‘developmental models’, ‘study models’, ‘toy models’, or ‘heuristic models’. The purpose of such model-systems is not to represent anything in nature; instead they are used to test and study theoretical tools that are later used to build representational models. In field theory, for instance, the so-called ϕ^4 -model has been studied extensively, but not because it represents anything in the world (it was well known right from the beginning that it does not), but because its simplicity allows physicist to study complicated techniques such as renormalization in a simple setting and get acquainted with mechanisms – in this case symmetry

Fifth, t-representation is not the only element in the above scheme whose absence is as interesting as its presence; structures and equations may similarly be construed. Although formalisations play an important role in modelling, not all scientific reasoning is tied to a formal apparatus. In fact, sometimes conclusions are established by solely considering a fictional scenario and without using formal tools at all. If this happens it is common to speak of a *thought experiment*. Although there does not seem to be a clear distinction between modelling and thought-experimenting in scientific practice, there has been little interaction between the respective philosophical debates.³⁷ This is lamentable because it seems to be important to understand how models and thought experiments relate to each other. In a recent paper Davies (2007) argues that there are important parallels between fictional narratives and thought experiments, and that exploring these parallels sheds light on many aspects of thought experiments. This take on thought experiments is congenial to the view on models presented in this paper and suggests that modelling and thought-experimenting are intrinsically related: thought experiments (at least in the sciences) are models without the formal apparatus.³⁸

Sixth, although Walton's general idea of rules of generation is intuitively clear, it turns out to be difficult to give an account of these rules. The two most important rules in the context of literary fiction – the reality principle and the mutual belief principle – suffer from intrinsic

breaking – which are important in other contexts (Hartmann 1995). It is advantage of the proposed view of modelling that it can account for this practice without problem.

³⁷ Extensive discussions of Thought experiments can be found in Brown (1991), Sorensen (1992), and Brown's and Norton's contributions on this topic to Hitchcock (2004).

³⁸ As an example consider Galileo's law of equal heights (Sorensen 1992, 8-9). Take a u-shaped cavity, put a ball on the edge on side, and let the ball roll down into the cavity. Galileo then argued that it would have to reach the same height at the other side – this is the law of equal heights. Of course Galileo realised that the ball's track was not perfectly smooth and that the ball faced air resistance, which is why the ball in an actual experiment does not reach equal height on the other side. So Galileo considered an idealised situation in which there are neither friction nor air resistance and argued that the law was valid in that scenario. This thought experiment fits the above account of model-systems: Galileo considered was a fictional scenario specified by a simple description, yet the conclusion he wanted to reach was not part of that description and was reached by using certain general principles that he took to be valid in situations like the one considered. Moreover, had Galileo used a mathematical machinery to derive his conclusion instead of informal arguments, physicists would refer to the product of his endeavour as a model. One would write down a curve specifying the shape of the cavity (for instance a parabola), specify its mechanical properties (frictionlessness), use mechanical laws to calculate the trajectory of the ball, and then find that it ends up at equal height on the other side. This is the sort of thing we find in mechanics textbooks, and which are referred to as mechanical models of a situation.

problems.³⁹ Worse, they may also lead to wrong results when put to work in science. So what are the rules of generation in scientific fictions? This is a substantial question that needs to be addressed, but we should not expect a single unified answer. On the contrary, it seems plausible to assume that different disciplines have different rules, and understanding what these rules are will shed light on how modelling in these disciplines works. So we should not expect a ready-made answer, but rather regard the study of rules of generation as part of research programme aiming at understanding the practice of modelling in various branches of science.

Seventh, not all models are introduced by verbal descriptions; sometimes we use drawings, sketches or diagrams to specify the model-system. There are linguistic and non-linguistic props. Although I have discussed pretence theory only in as far as it deals with linguistic props, the scope of the theory is much wider than that. In fact it covers all kinds of props, among them the classical media of visual art (paintings, drawings, etchings, etc.), as well as photography and film. So the current framework is equipped to deal with p-representation that is nonverbal.

Eighth, the fact the view of modelling advanced here is developed by drawing analogies with literary fiction should not be taken to suggest that there are no differences between the two. An in-depth comparison between literature and scientific modelling is beyond the scope of this essay, but some salient differences are readily stated. Literary plots are often complex and convoluted, while fictional scenarios canvassed in science are extremely simple and it seldom takes more than a few lines to describe the set-up. One of the reasons for this is that they must allow for mathematical treatment. Fictional scenarios in science are also often created with a specific target-system in mind, and the scenario is chosen such that t-representation can be set up – considerations that play only a marginal, if any, role in literature. Aesthetic considerations (style, genre, etc.) are irrelevant for model-descriptions, and so are emotional reactions of the reader to the plot. Finally, authorship is irrelevant in science: we often name models after their progenitors (e.g. the ‘Bohr model’), but this is merely a sociological fact

³⁹ Roughly, the Reality Principle says that if $p_1 \dots p_n$ are direct fictional truths, then proposition q is an indirect fictional truth iff: were it the case that $p_1 \dots p_n$, then it would be the case that q . The Mutual Belief Principle says that that if $p_1 \dots p_n$ are direct fictional truths, then proposition q is a indirect fictional truth iff: it is mutually believed in the artist’s society that were it the case that $p_1 \dots p_n$, it would be the case that q . See Walton (1990, Ch. 4) for a discussion of these principles.

with no systematic import since ambiguities and open questions are not resolved by appeal to the author's intention or context.

Ninth, needless to say, pretence theory is not without internal problems.⁴⁰ Although Walton's account eschews common-sense understandings of imagination (as noted above), more needs to be said about what exactly imagining amounts to in science and about how it differs from imagining in other contexts, as well as how it differs from other activities like considering, pondering, and entertaining. However, I will have to leave this issue for another occasion.

6. A First Stab at T-Representation

So far I have argued that models are imagined objects and I have shown how this leads to a coherent overall view of scientific modelling (shown in Figure 1). In particular, I have presented an account of what it means for claims about a model-system to be true, how we learn about model-systems, and how we can meaningfully compare them to either things in the world or other model systems. What is still missing from the analysis is an account of how model-systems represent (i.e. t-represent) something beyond themselves. The structuralist answer (that representation essentially is isomorphism) is not available to the fiction view since only structures can enter into isomorphisms and model-systems, on this view, are not structures. So we have to go back to the drawing board and develop a new account of representation that can explain how a model-system of the kind introduced in Section 4 can represent a target system. This is project for this section.

The first question is what to choose as our source from which we might formulate an account of t-representation. So far I have developed an account of scientific modelling by drawing analogies with literary fiction. Unfortunately this analogy does not seem to be productive when it comes to t-representation. Understanding t-representation involves establishing and understanding a relation between the fictional scenario and parts (or aspects) of the real world. While we sometimes do this casually (for instance when I compare my friend James with Zapp), there is controversy over whether this is in any way essential to our engagement with

⁴⁰ For critical discussion see, among others, Lamarque (1991), Budd (1992), and the contributions to the symposium on Walton's book in *Philosophy and Phenomenological Research* 51 (1991). See Currie (2004) for a discussion of difference notions of imagination.

fiction, and whether it leads to any interesting insights. Elgin (1996, Ch. 6) argues it does, which is what Kivy (2006, Chs. 24-28) denies. But even if this controversy could be resolved in favour of those who believe in the cognitive value of literature, there is no general and unambiguous method of bringing to bear literary fictions on real-world situations, which could serve as the blue-print for t-representation in science.⁴¹

The analogy I wish to exploit in what follows is the one between maps and scientific representations. This analogy is of course not new; see Sismondo and Chrisman (2001) for a survey and discussion. But I want to put the analogy to a slightly different use than other writers. While the map analogy has in the past mainly been employed to defend some sort of scientific realism, I wish to remain non-committal about realism and use maps only to explain how representation works at a the most general level.^{42, 43}

The essence of a map is that it allows us to ‘read off’ properties of the territory from the map: by looking at a map of London we see that Camden lies west of Hackney, Brixton is south of the river, etc. The map is different from a verbal description in that it does not merely state these facts; maps are not long lists with sentences describing a certain area. Facts about the city are inferred from facts about the map itself and a ‘key of translation’, which says how facts about the map translate into facts about the city. This is realisation provides us with the elements of the general scheme of representation:

X t-represents Y iff:

⁴¹ Elgin’s account is based on the notion of exemplification. This account is on the right track, and worked out version of the account I propose below will draw on many of its insights. However, at least in its basic form, this account does not cover cases in which the representational vehicle and the target do not share the relevant properties. The account suggested below is more permissive in that respect.

⁴² Throughout this chapter I use a realistic idiom in the sense that I assume that what is represented, the target system, exists. This is for the ease of formulation and my position could be restated from the point of view of *metaphysical* antirealism. What I want to remain non-committal about is *scientific* realism, roughly the position that theories are more or less truthful mirror images of reality. At a general level representing something 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. A general account of representation has to make room for non-realistic representations in this sense.

⁴³ Maps are of course real and not fictional objects. It will become clear as we proceed that representation works in the same way for fictional and real objects. Hence that maps, unlike model-systems, are material objects is no impediment to using them in the current context.

- (R1) X denotes Y .
- (R2) X comes with a key K specifying how facts about X are to be translated into claims about Y .

In nutshell, the idea is that the first condition establishes the aboutness of X , and the second guarantees the cognitive relevance of X for Y .⁴⁴ Before qualifying these conditions, let me illustrate them in more detail.⁴⁵ I have in front of me a map of North London. This is the first condition: the map denotes North London. Now I look at the details. I see a black rectangle on a black line and written next to it is ‘Camden Road’. The explanations that come with the map say that this rectangle stands for an over-ground railway station, the name next to it is the name of the station, and the black line stands for the rail tracks. A bit further up there is a black dot on a black line. The legend say that the dot stands for a tube station, and the name written next to it is the name of the station, in this case ‘Kentish Town’. Between the two there is a by thick yellow line, which stand for a main road. Hence, that a black rectangle labelled ‘Camden Road’ is connected with a thick yellow line to a black dot labelled ‘Kentish Town’ (a fact in the map) translates into the fact that Camden Road railway station is connected to Kentish Town tube station by a main road (a fact about North London). Furthermore, from the fact that this yellow line is 4.5cm long, I can infer that the actual distance between the two is about 1km since the scale of the map is 4.55cm to 1km. Finally, the ‘Kentish Town’ dot lies vertically above the ‘Camden Road’ rectangle, from which I infer that Kentish Town tube station is north of Camden Road railway station.

Our use of a map essentially involves a key, telling us how to translate facts about the map into facts about North London. Some elements of the key are stated at the bottom of the map; for instance, we are instructed that rectangles stand for railway stations and dots for tube stations. Other elements are conventions that are so common that they are assumed without further explanation. The top of the map indicates north, for example, and the distances in the map are proportional to distances in the world (where the ‘scale’ of the map gives the proportionality factor). But these are mere conventions and there is nothing ‘natural’ or ‘self

⁴⁴ The first condition is Goodman’s (1976, Ch. 1) who has argued that denotation lies at the heart of representation.

⁴⁵ Common alternatives to the current proposal are isomorphism and similarity accounts of representation; see Frigg (2006) and Suárez (2003) for discussions. Other alternatives have been proposed by Contessa (2007), Hughes (1997), Suárez (2004; 2006) and Toon (2009). For want of space I cannot discuss these here.

evident’ or even ‘necessary’ about them. We could use rectangles to denote tube stations rather than railway stations. We could draw the map so the south rather than north is on top, and have projection techniques that do not preserve distances.⁴⁶ The crucial point, though, is that what a map represents depends not only on facts *in* the map, but on the key that is used to translate these facts into claims about the world. And this key does not simply ‘jump off the page’; they are not ‘in’ the map itself. Instead, one has to know what the key is, and boundary-drawing diplomats had better take a class in map reading before attending international summits.

My claim is that model-systems are t-representations in the same way in which maps are: they denote a target system and certain facts obtain in them (in the sense explained in Sections 4 and 5) which are then translated into claims about a target system by using a key. As an example, consider the Bohr model of the hydrogen atom. On the current analysis this model consists of a model-system, which is specified by a model description and which is described by a formal apparatus (classical mechanics plus the Bohr-Sommerfeld quantisation rule). A number of facts obtain in the model-system, among them that it has discrete energy levels. We then take the model-system to denote real hydrogen atoms, and then use a simple key – here identity (more about this below) – to translate this fact into the claim that hydrogen itself has discrete energy levels.

Let me now add three qualifications. First, (R1) and (R2) provide the *general form* of an account of t-representation, which needs to be concretised in every particular instance of a t-representation. In fact, ‘denotation’ and ‘key’ are abstract in the sense introduced in Section 2 and need fitting out in every particular instance. In order to understand how a *particular* representation works, we need to account for how the particular *X* comes to denote the particular *Y*, and we have to provide a particular key *K*. In the above example, we borrowed denotation from ordinary language by saying ‘this is a map of North London’, and the key

⁴⁶ Nautical maps, for instance, use the Mercator projection system and do not preserve distances; they preserve angles and one obtains wrong results when translating the distance between two points on a map into the distance between two locations. And this mistake has been made over and over again. As Sismondo and Chrisman (2001, 42-43) point out, about half of a sample of 137 international maritime boundaries are not where they were meant to be. When diplomats met to draw the boundaries between territories they had these charts on the table. They intended draw the border half way between two territories and so they drew the line *on the map* mid-point between the territories. This is mistake: even relatively close to the equator the line thus drawn can be over 7km away from the actual line of equidistance.

was provided to us by cartography. But other cases may work differently since there may be different sources of denotation and there may be any number of keys that can be used to interpret X . Moreover, keys are often implicit and determined by context. This is often the case with scientific representations, which unlike maps, rarely, if ever, come with something like a legend. It is one of the challenges facing a philosophical analysis of representation to make hidden assumptions explicit, and present a clear statement of them. So there is much more to be said about t-representation than is contained in (R1) and (R2) – they are merely blanks to be filled in in every particular instance. Thus, the claim that something is a t-representation amounts to an invitation to spell out what how exactly X comes to denote Y and what K is.

Nonetheless, this generality is an advantage. The class of t-representations is large and its members varied. A view that claims that all t-representations work in exactly the same way would be doomed to failure right from the beginning. Maps, graphs, architectural plans, diagrams, photographs, (certain kinds of) paintings and drawings, and of course scientific models, are all t-representations in that they satisfy (R1) and (R2), but they work in very different ways. The differences between them are that these conditions are realised in very different ways: different keys are used and denotation has different sources. The challenge for a complete account of representation is to come up with a taxonomy of different ways in which the two conditions can be realised, and to explain how they differ from each other. Needless to say, this is a Herculean task that I cannot undertake here since there are many different kinds of keys. That said, the value of this account of representation is that it provides us with a framework in which to discuss these questions.⁴⁷

A second qualification I would wish to add to the scheme sketched above is to note that there is one important disanalogy between maps and scientific models: where their respective keys come from. In the case of the map we have the target system in front of us, we explore it *directly* (by taking measurements, etc.) and *then* we construct the map. So a map is an elegant summary of what someone already knows, and its sole purpose is to effectively summarise this knowledge and communicate it to those who are not in the business of land surveying.

⁴⁷ In passing I would like to point out that this account of representation satisfies the conditions of adequacy that I presented in my (2006). The *ontological puzzle* is addressed by the account of model-systems presented in Section 5. The *enigma of representation* is met by (R1) and (R2). The *problem of style* now becomes the question of how denotation works and what keys are used.

Science is not like this; we do not first survey the hydrogen atom and then construct a model to communicate the findings to those not yet familiar with it. We typically construct models to find out something genuinely new about the target system; something that *no one* yet knows.

This disanalogy does not undermine the saliency of t-representation for our analysis of modelling. Even if the process of constructing a model involves much more than elegantly summarising observations, once the model-system is constructed (no matter how!) we have to specify how it relates to the world, and this is done by providing a key. However, unlike for maps where we know the key by construction (we have used a certain projection method, certain symbols, etc. when drawing the map), in the case of models the key has the character of a hypothesis.⁴⁸ We stipulate that we expect the model to bear this or that relation to its target, and then evaluate this claim against the best available background knowledge and by subjecting it to test using the usual methods of scientific investigation. How exactly this is done depends on the details of the representation. That is, it depends on the key used and the nature of the denotation relation (for instance, an assessment of the accuracy of a key for model in elementary particle physics will be very different from the assessment of an engineering model of a bridge). Understanding these processes should be part of a future investigation into the nature of different kinds of t-representations (*cf.* the first qualification). For now it is sufficient to point out that keys can be hypothetical, and that this does not undermine the status of models as t-representations.

Third, (R2) states that we need a key specifying how to translate *facts* about *X* into *claims* about *Y*. This is not a slip. An acceptable definition of t-representation has to make room for misrepresentation. A map can contain errors in the sense that even if we use the right key and use it correctly we may obtain wrong results. For instance, it might have happened that the cartographers failed to connect the black dot and the black rectangle with a yellow line, and so we would have been led to believe that the two stations are not connected by a main road. This would not have turned the map into a non-t-representation; it would still have been a t-representation, but one that misrepresents North London. Saying that we translate facts about the map into *claims* about the target makes room for error because claims can be true or false,

⁴⁸ Although this is reminiscent of Giere's claim that models are connected to their target systems with a 'theoretical hypothesis' (1988, 80), the point is a different one. In Giere's account we call a claim to the effect that the model is similar to the target in specific way a theoretical hypothesis; the current view, by contrast, emphasises the hypothetical – fallible, tentative, and conjectural – character of keys attributed to a model.

while facts cannot. A representation is a *faithful* representation iff if all claims about *Y* are true.

There is now also a straightforward way to draw a delineation between cases of misrepresentation and cases of failure of representation. *X* is a misrepresentation if it is not faithful (and notice that misrepresentation comes in degrees!). Something is not a t-representation at all if either of the two conditions fails. We have a failure of (R1) if there is no target system; a map of Atlantis fails to be a t-representation of Atlantis because there is no Atlantis, and hence Atlantis cannot be denoted. By contrast, the failure can be put down on condition (R2) if *X* it has no intrinsic properties that are interpreted by using a key. This is why proper names, for instance, are not t-representations: they denote the bearer of the name, but there is no key that translates properties the *name itself* possess into claims about the bearer of the name. If, for some reason, one wants to call proper names ‘representations’ then one can do so, but it is important to realise that they are not t-representations, and being a t-representation is what matters both in the case of maps and in the case of scientific models.

With this in mind we can see what is wrong with Callender and Cohen’s (2006) argument that there is no special problem about scientific representation. Because scientific representation comes down to an act of arbitrary stipulation, by their reading, explaining how we make such stipulations lies in the province of philosophy of mind and not in the realm of philosophy of science at all. They ask: ‘Can the salt shaker on the dinner table represent Madagascar?’, and immediately reply ‘Of course it can, so long as you stipulate that the former represents the latter. [...] Can your left hand represent the Platonic form of beauty? Of course, so long as you stipulate that the former represents the latter.’ (*ibid*, 73-74). If all you mean by representation is denotation, then this is correct. But for something to be a t-representation, more than mere denotation is needed. We would need a key telling us how to translate certain properties of the salt-shaker into claims about Madagascar, or properties about my left hand into properties about the Platonic form of beauty, which, by their own admission, we don’t.⁴⁹

Why it is so important to be a t-representation, and why is simple stipulation not enough? The answer to this question is that maps as well as scientific representations belong to a category of representations that function cognitively: we study *X* to learn something about *Y* that we did not already know. In fact, model-systems are the units on which significant parts of

⁴⁹ For a more extensive discussion of Callender and Cohen’s argument see Toon (2009).

scientific investigation are carried out rather than on the target system 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. We do this by first finding out what is true in the model-system itself (*cf.* Sec. 5), and then translating the findings into claims about the target itself. This is possible only if the model-system is a t-representation in the above sense. Denotation is not enough for this to happen. Proper names don't inform us about the properties of things they stand for; we can turn and twist 'hydrogen' as long as we wish, but we won't thereby learn anything about hydrogen.

As I mentioned above, I regard the detailed study of different keys as a research programme to be undertaken in the future. However, to get a better idea of what such an investigation involves I now want to discuss two keys often used in science: identity and ideal limits. The simplest of all keys is *identity*, the rule according to which facts in the model (or at least a suitably defined class of facts) are also facts in the world. For example, if X t-represents Y by identity, then it follows from the fact that X has discrete energy levels that Y has discrete energy levels too. Although scientists often talk as if the relation between models and reality was identity, there are actually very few, if any, models that work in this way.

A more interesting key the *ideal limit* key. Many model-systems are idealisations of the target in one way or another. A common kind of idealisations is to 'push to the extreme' a property that a system possesses. This happens when we model particles as point masses, strings as massless, planets as spherical, and surfaces as frictionless. Two things are needed to render such idealisations benign: experimental refinements and convergence (Laymon 1991). First, there must be the possibility of *in principle* refining actual systems in a way that they are made to approach the postulated limit (that is, we don't actually have to produce these systems; what matters is that we in principle could produce them). With respect to friction, for instance, one has to find a series of experimental refinements that render a tabletop ever smoother and hence allow real systems to come ever closer to the ideal frictionless surface. These experimental refinements together constitute a sequence of systems that come ever closer to the ideal limit. 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 the behaviour in the 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 convergence. If there exists such a sequence of refinements and if the limit is monotonic, then the model is an ideal limit.

If a model is an ideal limit, this implies a key. To see how, let us first briefly recapitulate the mathematical definition of a limit. Consider a function $f(x)$, and then ask the question how $f(x)$ behaves if x approaches a particular value x_0 . We say that the number F is the limit of $f(x)$ (in symbols: $\lim_{x \rightarrow x_0} f(x) = F$) iff for every positive number ϵ (no matter how small), there exists another positive number δ such that: if $|x - x_0| < \delta$, then $|f(x) - F| < \epsilon$. Colloquially, this says that the closer x comes to x_0 , the closer $f(x)$ comes to F : if we know that x is less than δ way from x_0 , then we also know that $f(x)$ is less than ϵ away from F . This idea can now be used for ideal limits in the above sense. The sequence of experimental refinements plays the role of x , and the ideal limit itself is x_0 (in the example: the ever smoother table tops correspond to different values of x , and the frictionless plane corresponds to x_0). The behaviour of the object corresponds to f . If there is a limit we know that if the difference between the friction of the real plane and the ideal frictionless plane is smaller than δ , then difference between the behaviour of the real spinning top and the ideal spinning top in the model-system is smaller than ϵ . So if we are given the friction of the table, we know how to translate facts obtaining in the model-system into claims about the world.⁵⁰

Of course not all model-systems are ideal limits of their target-systems in this sense.⁵¹ For instance, we cannot possibly produce a sequence of systems in which Planck's constant approaches zero. In other cases it may not be clear whether there are such limits. For instance, mathematical knot theory is a branch of topology and as such it deals with one-dimensional

⁵⁰ I have smuggled in a premise here: that it makes sense to quantify differences in the friction of surfaces and the behaviour of spinning tops in terms of numbers. This is not implausible and could be made precise, for instance, by using friction coefficients and a geometrical measure for the closeness of trajectories. The following two questions are more pressing. First, how can we know whether or not a certain model-system is an ideal limit of the target at hand? Second, what is the relation between ϵ and δ ? In real applications one would like to know how close to the limit one would have to come to get a result that is precise to a particular degree. Typical mathematical existence results are of no help here. These are open questions that need to be addressed.

⁵¹ This corresponds to Rohrlich's distinction between factual and counterfactual limits (1989, 1165).

strings. But physical strings have finite width. Hence the question arises whether, and if yes, in what sense the results of mathematical knot theory carry over to physical situations. So it is an open question how to translate facts in idealised systems into claims about a real-world target if they are not ideal limits – or in the current idiom: there is a question about what they key is – one that should preoccupy us in the future.

7. Re-Reading the Newtonian Model of the Sun-Earth System

Case studies are the touchstone of philosophical analysis, and so it is imperative to show that the account developed in this chapter can shed light on typical cases of scientific modelling. For this reason I now discuss a standard example of a scientific model – the Newtonian model of the sun-earth system – and show that the fiction view not only has the resources to explain what happens in this case, but also makes features of the model visible that are usually overlooked. Hence, the fiction view of models, far from being an idle philosophical pastime, is actually a powerful tool to help us to better understand what is involved in scientific models.

The aim of the Newtonian model is to determine the orbit of the earth moving around the sun.⁵² We first posit that the only force relevant to the earth's motion is its gravitational interaction with the sun, and we neglect all other forces, most notably the gravitational interaction with the other planets in the solar system. This force is given by Newton's law of gravity, $F_g = Gm_p m_s / r^2$, where m_p and m_s are the masses of the earth and the sun respectively, r the distance between the two, and G the constant of gravitation. We then make the idealising assumption that both the sun and the earth are perfect spheres with a homogeneous mass distribution (i.e. the mass is evenly distributed over the sphere), which allows us to treat their gravitational interaction as if mass of both spheres was concentrated in their centre. The sun's mass is vastly larger than the earth's and so we assume that the sun is at rest and the earth orbits around it. Now we turn to classical mechanics and use Newton's equation of motion, $\vec{F} = m\vec{a}$, where \vec{a} is the acceleration of a particle, m its mass and \vec{F} the force acting on it. Placing the sun at the origin of the coordinate system and plugging in the

⁵² See, for instance, Feynman, Leighton, and Sands (1963, Secs. 9.7 and 13.4) and Young and Freedman (2000, Ch. 12).

above force law we obtain $\ddot{\vec{x}} = -Gm_s \vec{x} / |\vec{x}|^3$, the differential equation describing the earth's trajectory (where we have, of course, used $\vec{a} = \ddot{\vec{x}}$, i.e. that the acceleration is equal to the second derivative of the position). This equation can be solved and we find that the earth moves on an elliptic orbit around the sun.

When we read the above description, which tells us to regard the earth and the sun as ideal homogeneous spheres gravitationally interacting only with each other, this description serves as a prop and we engage in an authorised game of make believe. We imagine the entity described in the description, where the rules of direct generation are just the rules of ordinary English. We understand the terms occurring in the description and we imagine an entity which has all the properties that the description specifies. The result of this process is the *model-system*, the fictional scenario which is the vehicle of our reasoning: an imagined entity consisting of two spheres, etc. The part of the above description that prescribes us to imagine the model-system is the *model-description*. Now focus on the formal apparatus.

$\ddot{\vec{x}} = -Gm_s \vec{x} / |\vec{x}|^3$ is the *model-equation*, which, in this case, is obtained from a general theory – Newtonian mechanics – by specifying the number of particles and their interaction. This equation specifies a *model-structure*, which is instantiated in the model-system (*cf.* Section 6). A proper analysis of the structure described by this equation would require formal techniques that are beyond chapter.⁵³ But for our purposes nothing hinges on giving all the details (since our concern here is not the applicability of mathematics); what matters at this point is only that such analysis can be given and that its upshot is that the model-equation applies to the model-system (and is literally true of it). The model-equation then is the formal expression of a principle of indirect generation. Using this principle we find that it is true in the model-system that the light sphere orbits around the heavy sphere on an elliptical orbit. This is an implied truth because it has not been written into the model-description; it is something that we infer from the basic features of the model-system (as given by the model-description) and the rule of generation.

The next step is to connect our model to the target-system. We find clues about how to do this in the above description. Right at the beginning we are told that the model we are constructing is a model of the sun-earth system. This establishes denotation, which is condition (R1). As in

⁵³ Such an analysis can be found in Balzer, Moulines, and Sneed (1987, 29-34, 103-108, 180-191), Frigg (2003, Ch. 8), and Muller (1998, 259-66).

the above examples, we borrow denotation from ordinary language by using the expressions ‘sun’ and ‘earth’, which we take to refer to the relevant heavenly bodies. Should these expressions for some reason fail to refer, then t-representation would fail too. Ordinary language also plays a role in specifying the key. The first element of the key is the definition of an object-to-object correlation: we say that the light sphere in the model-system corresponds to the earth and the heavy one to the sun. Now things get more involved. We have made several idealisations (that the sun and the earth are spherical, that there are no forces other than the gravitational interaction of sun and earth, etc.) and we now have to say how these should be understood. Unfortunately physics texts usually do not say much about this question, or remain altogether silent about it. So at this point we have to appeal to philosophical theories of idealisation and the keys they imply. On a plausible reading of the Newtonian model, the idealisations made are taken to be ideal limits in the sense discussed in the last section. The limit is complex and involves many properties, but the leading idea is that we could – in principle – produce a sequence of systems where the forces acting on the sun and the earth become increasingly smaller and eventually converges towards zero (which would be done by taking more and more matter out of the universe). We can then also – again, in principle – produce a sequence of sun-earth systems in which the sun and the earth become ever rounder and their mass distributions ever more homogeneous. The claim then is that, first, in the limit the sequence of these systems converges towards the model-system (which is true by construction); second, the behaviour of the systems in this sequence converges towards the behaviour of the model-system (this is the ideal limit). Given this, we know how to translate claims about the model into claims about the target: if the actual target is less than δ away from the model-system, then the behaviour of the actual target is less than ε away from the behaviour of the model-system. This is (R2).

Asserting convergence between sequence and system constitutes a substantial claim that does not follow from the construction of the sequence. In fact, we have cannot strictly *prove* that this is so. This illustrates the hypothetical character of keys: they are postulated as a hypothesis and not given to us as in the case of the map. However, this does not mean that any hypothesis is as good as any other. We justify the stipulation of the ideal limit key (rather than another key) in two ways. First we appeal to background knowledge: we have tested the law of gravity and Newton’s equation of motion in countless situations and have good reasons to assume that it provides true descriptions in scenarios like the model-system. We derive predictions from the model-system (the trajectory of the earth) and compare them with

observations. At this point the ideal limit key becomes essential. If we have in ideal limit, then we know how the behaviour of the model-system relates to the behaviour of the target. Assume now we can sensibly quantify such distances (*cf.* footnote 50) and, given what we know about the universe, the forces and masses are such that the actual target-system is less than δ away from the model-system, then we can compare the theoretical trajectories of the earth with the observed ones and see whether they are less than ϵ away from each other. If this is the case, then this confirms our hypothesis that the model-system is an ideal limit. But notice – to come back to the point made in Section 2 – that what the model-system represents is not data, nor is there anything in the model that is directly comparable to data. The data used to confirm the model are obtained with the aid of specific observational techniques (optical telescopes, radio telescopes, etc.) and the character of the data varies with these techniques. *Given* a particular technique (and the theories behind it), the model can be used to calculate what one would have to observe; but the result of this calculation is not in any way part of the make-up of the model.

With all this in place, we can then start translating facts about the model-system into claims about the world. For instance, calculations reveal that the model-earth move on ellipses, and given that the model-system is an ideal limit of the target we can infer that real earth move on trajectories that are almost ellipses (or more precisely, on trajectories that are not more than ϵ away from an ellipse).

This is a complete analysis of the model of the sun-earth system. Hence, we see that the fiction view of models is able to provide us with a complete account of how scientific models work, and it can do so without having to go at great length to reconstruct scientific practice in terms of a particular revisionary philosophy (such as the structuralist programme). First appearances notwithstanding, the fiction view of models is close to scientific practice and provides an analysis of modelling that scientists would recognise. The fiction view of models, then, is an account of scientific modelling that is both philosophically well founded and close to scientific practice –the kind of account of modelling that we have been looking for.

10. Conclusion

I have argued that scientific modelling shares important aspects in common with literary fiction, and that therefore theories of fiction can be brought to bear on issues in connection with modelling. I have identified six such issues and suggested that pretence theory offers satisfactory responses to them. From this discussion emerges a general picture of scientific modelling, which views scientific modelling as a complex activity involving the elements shown in Figure 1. I have then used the analogy with maps to present the broad outlines of an account of t-representation and have shown how this account can be used to analyse how a typical model in physics, the Newtonian model of the sun-earth system, represents.

Acknowledgements

I would like to thank José Díez, Matthew Hunter, and Julian Reiss for helpful discussions and comments on earlier drafts.

Bibliography

- Balzer, Wolfgang, C. Ulises Moulines, and Joseph D. Sneed (1987), *An Architectonic for Science: The Structuralist Program*. Dordrecht.
- Bogen, James, and James Woodward (1988), "Saving the Phenomena", *Philosophical Review* 97:303-352.
- Boolos, George S., and Richard C. Jeffrey (1989), *Computability and Logic*. 3rd ed. Cambridge: Cambridge University Press.
- Brown, James (1991), *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*. London: Routledge.
- Budd, Malcolm (1992), "Review of 'Mimesis as Make-Believe'", *Mind* 101:195-198.
- Callender, Craig, and Jonathan Cohen (2006), "There Is No Special Problem About Scientific Representation", *Theoria* 55:7-25.
- Campbell, Norman (1920), *Physics: The Elements*. Cambridge: Cambridge University Press. (Reprinted as *Foundations of Science*. New York: Dover 1957.).

- Carnap, Rudolph (1938), "Foundations of Logic and Mathematics", in Otto Neurath, Charles Morris and Rudolph Carnap (eds.), *International Encyclopaedia of Unified Science. Vol. 1.*, Chicago: University of Chicago Press, 139-213.
- Cartwright, Nancy (1983), *How the Laws of Physics Lie*. Oxford: Oxford University Press.
- (1999), *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.
- Contessa, Gabriele (2007), "Scientific Representation, Interpretation, and Surrogate Reasoning", *Philosophy of Science* 74 (1):48–68.
- Crittenden, Charles (1991), *Unreality: The Metaphysics of Fictional Objects*. Ithaca and London: Cornell University Press.
- Currie, Gregory (1990), *The Nature of Fiction*. Cambridge: Cambridge University Press.
- (2004), "Imagination and Make-Believe", in Berys Gaut and Dominic McIver Lopes (eds.), *The Routledge Companion to Aesthetics*, London: Routledge, 335-346.
- Da Costa, Newton, and Steven French (1990), "The Model-Theoretic Approach to the philosophy of Science", *Philosophy of Science* 57:248-265.
- Davies, David (2007), "Thought Experiments and Fictional Narratives", *Croatian Journal of Philosophy* 7 (19):29-45.
- Downes, Stephen (1992), "The Importance of Models in Theorizing: A Deflationary Semantic View", *Philosophy of Science (Proceedings)* Vol. 1:142-153.
- Duhem, Pierre (1906), *La Théorie Physique, son Objet et sa Structure*. 2nd ed., Paris 1914 (Engl. trans. by Philip P. Wiener: *The aim and Structure of Physical Theory*. Princeton 1954).
- Elgin, Catherine Z. (1996), *Considered Judgment*. Princeton: Princeton University Press.
- Evans, Gareth (1982), *The Varieties of Reference*. Edited by John McDowell. Oxford: Oxford University Press.
- Feynman, Richard, Robert B. Leighton, and Mathew L. Sands, eds. (1963), *The Feynman Lectures in Physics*. Redwood City and Reading/Ma 1989.
- Fine, Arthur (1993), "Fictionalism", *Midwest Studies in Philosophy* 18:1-18.
- French, Steven (1999), "Models and Mathematics in Physics: The Role of Group Theory", in Jeremy Butterfield and Constantin Pagonis (eds.), *From Physics to Philosophy*, Cambridge: Cambridge University Press.
- French, Steven, and James Ladyman (1997), "Reinflating the semantic approach", *International Studies in the Philosophy of Science* 13:103-121.
- Friend, Stacie (2007), "Fictional Characters", *Philosophy Compass* 2 (2):141–156.

- Frigg, Roman (2003), *Re-presenting Scientific Representation*, PhD Thesis. London: University of London.
- (2006), "Scientific Representation and the Semantic View of Theories", *Theoria* 55:49-65.
- (2009), "Models and Fiction", *Synthese* (forthcoming).
- Giere, Ronald N. (1988), *Explaining Science: A Cognitive Approach*. Chicago: Chicago University Press.
- (2004), "How Models Are Used to Represent Reality", *Philosophy of Science* 71 (4):742–752.
- Godfrey-Smith, Peter (2006), "The Strategy of Model-Based Science", *Biology and Philosophy* 21:725-740.
- Goodman, Nelson (1976), *Languages of Art*. 2nd ed., Indianapolis and Cambridge: Hackett.
- Grüne-Yanoff, Till, and Paul Schweinzer (2008), "The Roles of Stories in Applying Game Theory", *Journal of Economic Methodology* 15 (2):131-146.
- Hacking, Ian (1983), *Representing and Intervening*. Cambridge: Cambridge University Press.
- Hartmann, Stephan (1995), "Models as a Tool for Theory Construction: Some Strategies of Preliminary Physics", in William E. Herfel, Wladiyslaw Krajewski, Ilkka Niiniluoto and Ryszard Wojcicki (eds.), *Theories and Models in Scientific Processes (Poznan Studies in the Philosophy of Science and the Humanities 44)*, 49-67.
- (1999), "Models and Stories in Hadron Physics", in Mary Morgan and Margaret Morrison (eds.), *Models as Mediators. Perspectives on natural and social science*, Cambridge: Cambridge University Press, 326-346.
- Hempel, Carl G. (1965), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.
- Hesse, Mary (1963), *Models and Analogies in Science*. London: Sheed and Ward.
- Hitchcock, Christopher, ed. (2004), *Contemporary Debates in Philosophy of Science*. Oxford: Blackwell.
- Howell, Robert (1979), "Fictional Objects: How They Are and How They Aren't", *Poetics* 8:129-177.
- Hughes, R. I. G. (1997), "Models and Representation", *Philosophy of Science* Vol. 64, Supplement:S325-S336.
- Kivy, Peter (2006), *The Performance of Reading: An Essay in the Philosophy of Literature*. Oxford: Blackwell.

- Lamarque, Peter (1991), "Essay Review of 'Mimesis as Make-Believe: On the Foundations of the Representational Arts' by Kendall Walton", *Journal of Aesthetics and Art Criticism* 49 (2):161-166.
- Lamarque, Peter, and Stein Haugom Olsen (1994), *Truth, Fiction, and Literature*. Oxford: Clarendon Press.
- Laymon, Ronald (1991), "Thought Experiments by Stevin, Mach and Gouy: Thought Experiments as Ideal Limits and as Semantic Domains", in Tamara Horowitz and Gerald J. Massey (eds.), *Thought Experiments in Science and Philosophy*, Savage, MD: Rowman and Littlefield, 167-191.
- Lewis, David (1978), "Truth in Fiction", in David Lewis (ed.), *Philosophical Papers, Volume I*, Oxford: Oxford University Press 1983, 261-280.
- McAllister, James W. (1997), "Phenomena and Patterns in Data Sets", *Erkenntnis* 47:217-228.
- McCloskey, Donald N. (1990), "Storytelling in Economics", in Christopher Nash (ed.), *Narrative in Culture: The Uses of Storytelling in the Sciences, Philosophy, and Literature*, London: Routledge, 5-22.
- Morgan, Mary (2001), "Models, Stories and the Economic World", *Journal of Economic Methodology* 8 (3):361-384.
- (2004), "Imagination and Imaging in Model Building", *Philosophy of Science* 71 (4):753–766.
- Morgan, Mary, and Margaret Morrison, eds. (1999), *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.
- Muller, Frederik A. (1998), *Structures for Everyone*. Amsterdam.
- Parsons, Terrence (1980), *Nonexistent Objects*. New Haven: Yale University Press.
- Rohrlich, Fritz (1989), "The Logic of Reduction: The Case of Gravitation", *Foundations of Physics* 19:1151-1170.
- Russell, Bertrand (1919), *Introduction to Mathematical Philosophy*. London and New York: Routledge 1993.
- Russell, Bertrand (1905), "On Denoting", in, *Logic and knowledge*, London Routledge 1956, 39-56.
- Shapiro, Stewart (2000), *Thinking About Mathematics*. Oxford.
- Sismondo, Sergio, and Nicholas Chrisman (2001), "Deflationary Metaphysics and the Nature of Maps", *Philosophy of Science (Proceedings)* 68 (38-49).

- Sklar, Lawrence (2000), *Theory and Truth. Philosophical Critique Within Foundational Science*. Oxford: Oxford University Press.
- Smith, Leonard (2007), *Chaos. A Very Short Introduction*. Oxford: Oxford University Press.
- Smolin, Lee (2007), *The Trouble with Physics: The rise of String Theory, the Fall of a Science, and What Comes Next*. London: Allen Lane.
- Sorensen, Roy (1992), *Thought Experiments*. New York: Oxford University Press.
- Suárez, Mauricio (2003), "Scientific Representation: Against Similarity and Isomorphism", *International Studies in the Philosophy of Science* 17 (3):225-244.
- (2004), "An Inferential Conception of Scientific Representation", *Philosophy of Science (Supplement)* 71:767-779.
- Suárez, Mauricio, and Albert Solé (2006), "On the Analogy between Cognitive Representation and Truth", *Theoria* 55:39-48.
- Sugden, Robert (2000), "Credible Worlds: The Status of Theoretical Models in Economics", *Journal of Economic Methodology* 7 (1):1–31.
- Suppe, Frederick, ed. (1977), *The Structure of Scientific Theories*. Urbana and Chicago.
- Suppes, Patrick (1960), "A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences", in Patrick Suppes (ed.), *Studies in the Methodology and Foundations of Science: Selected Papers from 1951 to 1969*, Dordrecht Reidel 1969, 10-23.
- Teller, Paul (2001), "Twilight of the Perfect Model Model", *Erkenntnis* 55:393–415.
- (2001), "Whither Constructive Empiricism", *Philosophical Studies* 106:123-150.
- Thomson-Jones, Martin (2007), "Missing Systems and the Face Value Practice", *Synthese* forthcoming.
- Toon, Adam (2009), "Models as Make-Believe", in Roman Frigg and Matthew Hunter (eds.), *Beyond Mimesis and Convention: Representation in Art and Science*, Berlin Springer.
- Vaihinger, Hans (1911), *The Philosophy of 'As If': A System of the Theoretical, Practical, and Religious Fictions of Mankind*. English Translation London: Kegan Paul 1924.
- van Fraassen, Bas C. (1980), *The Scientific Image*. Oxford: Oxford University Press.
- (1989), *Laws and Symmetry*. Oxford: Clarendon Press.
- (1997), "Structure and Perspective: Philosophical Perplexity and Paradox", in Marisa L. Dalla Chiara (ed.), *Logic and Scientific Methods*, Dordrecht: Kluwer, 511-530.
- (2002), *The Empirical Stance*. New Haven and London.
- Walton, Kendal L. (1990), *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge/MA: Harvard University Press.

Woodward, James (1989), "Data and Phenomena", *Synthese* 79:393-472.

Young, Hugh D., and Roger Freedman (2000), *University Physics with Modern Physics*. 10th ed. San Francisco and Reading (MA): Addison Wesley.