Of Rabbits and Men

Published in Bradley Armour-Garb and Fred Kroon (eds):
Fictionalism in Philosophy, New York: Oxford University Press 2020, 187-206.

Roman Frigg and Fiora Salis

1. Introduction

Imagine you get a pair of new-born rabbits, one male and one female, which you keep in your garden. After six months they mate and after another six months the female gives birth to a new pair of rabbits, again a female and a male. At this point you start wondering how your rabbit population will evolve over time. In the absence of crystal balls and fortune-tellers, your best bet is to construct a model of the situation. You begin by making some modelling assumptions. You assume that your rabbits always mate six months after birth and that the female of each pair gives birth to exactly one male-female pair another six months after mating. To keep your calculations manageable, you further assume that the rabbits don't die and that there are no restrictions on food supplies and living space. Let us label the times when rabbits mate and give birth by $t_1, t_2, \ldots$, where t_1 is the moment when you get the first pair of rabbits and two consecutive instants are always separated by a six-months interval; furthermore let $N(t_i)$ be the number of rabbit pairs at a certain instant of time t_i. Under these assumptions, some arithmetic shows that that the number of rabbit pairs at a given time t_i is the sum of the numbers of pairs at the previous two times: $N(t_i) = N(t_{i-1}) + N(t_{i-2})$. For instance, the rabbit population at time t_3 is $N(t_3) = N(t_2) + N(t_1) = 1 + 1 = 2$ pairs. Using this formula one quickly finds the rabbit pair numbers at all future moments: 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, … These are known as the Fibonacci numbers, named after Leonardo of Pisa whose nickname was 'Fibonacci'. He presented the population model we have just seen in his *Liber Abaci* which was published in 1202.

If you now figure out that after only five years you will already have 89 rabbit pairs and you get all excited about your success as a breeder, you may want to take a deep breath and look at your calculations again. What you have constructed is a model of the rabbit population. Your model world is one where rabbits procreate at fixed instants of time, where every female gives birth to exactly one male-female pair, where rabbits never die, and where no limitations on food and space constrain population growth. Real rabbits are not like that, and their living environment doesn't match the conditions of the model. So the calculations of the model are probably roughly correct for the first few time steps, after which model and real-world rabbit numbers will start diverging. There is no problem with this. No one will be foolish enough to think that they have immortal rabbits in their infinite garden. The scenario underlying the model is a *fictional* scenario, and model users know this.

The Fibonacci population model is no exception. Scientific discourse is rife with passages that *appear* to be descriptions of systems in a particular discipline, and the pages of textbooks and journals are filled with discussions of the behaviour of those systems. Students of mechanics investigate the dynamical properties of a point mass hanging from the ceiling on a massless string when studying the motion of a pendulum; 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. As in Fibonacci's case, the entities described in the model don't exist in the real world. Thomson-Jones (2010, p. 284) refers to such descriptions as 'descriptions of a missing system'. These descriptions are embedded in what he calls the 'face value practice': the practice of talking and thinking about these systems as if they were real systems. We observe that the size of the rabbit population grows monotonically in much the same way in which we say that the mass of the moon is approximately 7.34×10^{22}kg. Yet the former statement is about something that doesn't exist, while the latter is about a real object.

The face-value practice raises a number of questions. What account should be given of these descriptions and what sort of objects, if any, do they describe? The fiction view of modelling advances the thesis that these questions are best answered by drawing a parallel between scientific models and literary fiction. Peter Godfrey-Smith offers the following programmatic statement of the view:

> '[…] I take at face value the fact that modelers often *take* themselves to be describing imaginary biological populations, imaginary neural networks, or imaginary economies. […] Although these imagined entities are puzzling, I suggest that at least much of the time they might be treated as similar to something that we are all familiar with, the imagined objects of literary fiction. Here I have in mind entities like Sherlock Holmes' London, and Tolkien's Middle Earth. […] the model systems of science often work similarly to these familiar fictions.' (2006, p. 735)

This is the core of the fiction view of models: scientific models are akin to places and characters in literary fiction.

However, fictional discourse and fictional entities are beset with well-known difficulties, and hence explaining model-systems in terms of fictional characters and places seems to amount to little more than to explaining *obscurum per obscurius*. The challenge for proponents of the fiction view is to show that drawing an analogy between models and fiction has heuristic value. Such an analysis is missing from the literature published before 2010, and filling this gap was the project of a number of recent publications. In this essay present our own answer. Alternative approaches have been developed, among others, by Contessa (2010), Levy (2015), Thomasson (forthcoming), Thomson-Jones (Thomson-Jones forthcoming) and Toon (2012). Space constraints prevent us from engaging with these positions.

2. Models

In this section we formulate an account of fiction, which, we claim, explains what scientific models are.[1] In the next section we present an account of how they represent real-world target systems.

[1] This section is based on Figg (2010) and Salis and Frigg (forthcoming).

Before delving into a discussion of what kind of fictions models are, we ought to set the criteria by which we evaluate proposals. The most important selection criteria are what we call 'truth in fiction' and 'learning about fiction'. There is right and wrong in a discourse about model-systems just as there is right and wrong in literary discourse. It's true that Sherlock Holmes is a detective; it's wrong that he's a ballet dancer. It's true that the population in Fibonacci's model never decreases; it's wrong that the population size oscillates. But on what basis are such claims qualified as true or false, in particular if the claims concern issues about which the original description of the system remains silent? We need an account of truth in fiction, 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. And *nota bene* that this requirement concerns truth *within* a model. Whether what is true in a model is also true of the model's target system is an altogether different issue. It's true in Fibonacci's model that the population grows monotonically; it's false that the population of rabbits in our garden grows in this way. We turn to the model-world relation in the next section.

The second requirement is that we need an account of how we come to *know* truths about models. Scientists investigate models and find out about them. In fact, scientists engage with models precisely because they want to explore their properties. How do they do this?

The philosopher of science may now turn to the rich literature on fiction that has grown in aesthetics and metaphysics. Fictional anti-realists hold that there are no fictional entities and offer analyses of our discourse about them that don't incur ontological commitments (Everett 2013; Kroon 2011; Walton 1990). Fictional realists claim that there are fictional entities but they disagree on what sort of entities they are. Neo-Meinongians think that they are concrete non-existent objects (Parsons 1980), or possible but non-actual objects (Berto 2011). Abstract object theorists submit that they are abstract eternal Platonic entities (Zalta 1988), or abstract artefacts akin to other social constructs (Thomasson 1999). Giere (1988) and Weisberg (2013) defended abstract object views of models. Thomasson (forthcoming) and Thomson-Jones (forthcoming) develop an abstract artefact view of models. Barberousse and Ludwig (2009), Frigg (2010), Toon (2012) and Levy

(Levy 2015) proposed anti-realist accounts. A discussion of the pros and cons of these views is beyond the scope of this essay.[2]

Our own preference is for an account of modelling based on pretence theory, and we take our lead from Walton (1990). The point of departure of this theory is the capacity of humans to imagine things.[3] 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 a 'prop'.[4] '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 'principle of generation', prescribing what is to be imagined as a function of the presence of the object. If someone imagines something because she is encouraged to do so by the presence of a prop she is engaged in a game of make-believe. Someone who is involved in a game of make-believe is pretending; so 'pretence' refers to the participation in such a game and has (in this context) nothing to do with deception. The simplest examples of games of make-believe are cases of child's play. 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.

Pretence theory considers a vast variety of different props ranging from written text to movies, from paintings to plays, and from music to children's games. In the present context we only discuss the case of literature. Works of literary fiction are, on the current account, props because 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.

[2] See Salis (2016) for a discussion of fictional realism and anti-realism in the context of scientific modelling.

[3] Imagination can be propositional and does not carry any commitment to mental imagery. See Salis and Frigg (forthcoming) for a discussion of imagination.

[4] In ordinary English the term 'prop' refers to an object used by an actor on a theatre stage or film set during the performance. In Walton's pretence theory 'prop' has become a technical term referring to an object that has a rule attached to it that requires spectators to imagine a particular thing when they see the object.

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

According to Walton, 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 to something beyond themselves; representations are things that possess the social function of serving as props in authorised games of make-believe. We 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'. 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.

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. What rules can legitimately be used to reach conclusions of this sort is a difficult issue fraught with controversy. We will encounter examples of scientific principles of generation shortly, which will illustrate how rules of generation work. 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 models. Typically, models 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’, ‘assume’ or ‘imagine’ and thereby make it clear that they are not descriptions of fact, but an invitation to 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 reality or truth. 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.

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 population grows monotonically from the basic assumptions of Fibonacci’s model and some basic facts of arithmetic. This is explained naturally in the idiom of pretence theory. What is explicitly stated in a model description (that the rabbits breed in six month intervals, 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.

What exactly do we assert when we qualify 'Zapp drives a convertible' as true in the fiction while 'Zapp drives a lorry' as false?[5] We are not meant to believe statements made in a fiction, 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*). Truth and truth in fiction are distinct; in fact, truth in fiction is best not regarded as a species of truth at all. 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'.

The question now becomes: when is a proposition *p* fictional in a work of fiction *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 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, imaginings 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 everybody involved in the game has the same imaginings.

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*, and whether a proposition is to be

[5] In what follows we only deal with what has been called intrafictional statements. We set aside metafictional and transfictional statements, which require a different treatment. For a discussion of these see Salis (2016).

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, 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.

This analysis of truth in fiction carries over to models by replacing *p* by a claim about the model, *w* by the description of the model, and *w*-principles of generation by the laws and principles assumed to be at work in the model. For instance, 'the population grows monotonically' is true in Fibonacci's population model iff the description of the system together with the laws and principles assumed to hold in the system (namely the laws of arithmetic) imply that this is the case.

We now also see how to answer the epistemic question of how we learn about models: 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 because a good deal of the work that scientists do with models can accurately be described as studying consequences of the basic model assumptions.

With this in place, we can now distinguish two different notions of representation. As noted previously, 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. However, the term is used rather differently in both science and philosophy of science where it denotes a relation between the model and its target. These two notions of representation are complementary – we will turn to this point in the next section. For now it is important not to get them mixed up, and for this reason we call the former 'p-representation' ('p' for 'prop') and the latter 't-representation' ('t' for target). So p-representation is the ability of prop to prescribe imaginings; t-representation is the relation between a model and a real-world target system. In this terminology pretence theory can be understood as an analysis of p-representation. This leaves pending an analysis of t-representation, to which we turn in the next section.

3. Representation

Models often represent selected aspects or parts of the real world. We refer to such a part or aspect as a model's *target system*. The intended target is often mentioned in the name of the model, for instance when we speak of the ball and spring model *of a polymer* and the Bohr model *of the atom*. Labels like these identify polymers and atoms as target systems. The relation between the model and its target system is t-representation.

To develop an account of t-representation we take as our point of departure Nelson Goodman and Catherine Elgin's concept of *representation-as*. We first introduce this concept as developed by Goodman and Elgin and then modify it so that it can be applied to scientific models.[6] Many instances of representation represent their target as being thus and so. Holbein's *Portrait of Henry VIII* represents Henry as imposing and powerful, and a caricature of Churchill represents him as a bulldog. The leading idea is that scientific representation works much the same way.

To analyse 'representation as' we first have to put few notions in place. The first is the notion of representation *of*. One of Goodman's central posits is that denotation is 'the core of representation' (1976, p. 5). Holbein's portrait denotes Henry VIII and the Fibonacci model denotes a population of rabbits. In that sense the painting and the model are representations *of* their respective targets. To distinguish being a representation of something from other notions of representation we introduce the locution 'representation-of', and to have terminological unanimity we also speak of 'representation-as'. Denotation is what establishes representation-of; but as we will see it is not sufficient to establish representation-as.

Not all representations are a representation-of. A picture showing a unicorn is not a representation-of a unicorn because things that don't exist can't be denoted. Yet there

[6] This section is based on Frigg and Nguyen (2016; 2018); for an extensive review of different theories of representation see their (2017).

is a sense in which such a picture is a representation – after all it *shows* a unicorn. Goodman and Elgin's solution to this is to distinguish between being a representation-of something and being a something-representation (Elgin 2010, pp. 1-2; Goodman 1976, pp. 21-26). A picture showing a unicorn is a unicorn-representation but it is *not* a representation-of a unicorn. In general, an object *X* is a *Z*-representation if it portrays *Z*. The crucial point is that this does not presuppose that *X* is a representation-of *Z*; indeed *X* can be *Z*-representation without denoting anything. A picture, for instance, must denote a man to be a representation-of a man. But it need not denote anything to be a man-representation. This raises the question of what it takes for *X* to be a *Z*-representation. For pictures this is a much-discussed question. Perceptual accounts hold that picture *X* is a Z-representation if, under normal conditions, an observer sees a *Z* in *X* (Lopes 2004). Goodman and Elgin analyse Z-representation in terms of *X* belonging to the genre of *Z*-representations (Elgin 2010, 2-3; Goodman 1976, 23). Neither of these suggestions is useful when dealing with scientific models. The objects that constitute models (immortal rabbits and the like) neither make observers see a *Z* in them, nor do they belong to genres in any obvious way. What turns them into *Z*-representations is an act of interpretation by a user: we *interpret* the objects that constitute *X* in terms of *Z*. A fictional collection of immortal rabbits is a population-representation because we interpret them as population. This may seem natural, but interpretations are not always that obvious. A collection of billiard balls becomes a gas-representation when we interpret billiard balls as gas molecules (as in Boltzmann's ideal gas model); a checkerboard becomes as social-segregation-representation when we interpret the squares as locations in a city (as in Schelling's segregation model); and a system of pipes and reservoirs becomes an economy-representation when we interpret the reservoirs as economic agents like banks and the flow of water as the flow of commodities.

Next in line is *exemplification*. An item exemplifies a property if it at once instantiates the property and refers to it. As Goodman puts it: 'Exemplification is possession plus reference. To have without symbolising is merely to possess, while to symbolise without having is to refer in some other way than by exemplifying' (Goodman 1976, p. 53). Paradigmatic examples of this are samples. A chip of paint on a manufacturer's sample card both instantiates a certain colour, and at the same time refers to that colour (Elgin 1983, p. 71).

Even though exemplification requires instantiation, not every property instantiated by an object is exemplified by it. The chip of paint does not, for example, exemplify its shape or its location with respect to the other samples. In order to exemplify a property, an object must both instantiate the property and the property itself must be made epistemically salient. How saliency is established will be determined on a case-by-case basis, depending on context and epistemic interest.

We now have the tools to analyse representation-as. The structure of the locution is that an object *X* represents a subject *T* as being *Z*. A first stab would be to say that *X* represents *T* as *Z* if *X* is a *Z*-representation and denotes *T*. This is on the right track, but one crucial element is missing. It is important that at least some properties of *Z* are ascribed to *X*. A caricature portraying Churchill as a bulldog misses the point if it does not manage to impute a number of bulldog-properties to Churchill: the caricature portrays him as aggressive, bullish, and intimidating. Representation-as can then be defined as follows: *X* represents *T* as *Z* if, and only if, (i) *X* denotes *T*, (ii) *X* is a *Z*-representation exemplifying certain properties, and (iii) *X* imputes these properties, or related ones, onto *T* (Elgin 2010, p. 10).

Our core claim is that t-representation is representation-as. But to make this idea tick, the third condition needs some qualification (Frigg and Nguyen 2018). The definition somewhat vaguely says that *X* imputes certain properties, *or related* ones, onto *T*. The motivation for adding this clause is that the properties exemplified by a scientific model and the properties imputed to its target system need not be identical. In fact, few models portray their targets as exhibiting exactly the same features as the model itself. Fibonacci's rabbits instantiate immortality, but when used as a model of a concrete population of flesh and blood animals, immortality is not imputed onto them. What is imputed is the related property of living long enough for the initial few generations to develop without the interference of mortality (or some such property). The problem with invoking 'related' properties is not its correctness, but its lack of specificity. Any property can be related to any other property in some way or other, and as long as no specific relation is specified it remains unclear which properties are imputed onto the system.

For this reason it is preferable to build a specification of the relationship between model properties and target properties directly into an account of t-representation. We call such a specification a *key*. A key in effect translates one set of properties (the ones exemplified by the model) into another set of properties (the ones imputed to the target). This key can but need not be identity; any rule that associates a unique set of imputation properties with the model properties is admissible. The relevant clause in the definition of representation-as then becomes: *X* exemplifies one set of properties and imputes another set of properties to *T* where the two sets of properties are connected to each other by a key.

Maps furnish a simple example. The model property is the measured distance on the map between the point labelled 'Rome' and the point labelled 'London'; the imputation property is the distance between Rome and London; and the key is the scale of the map. So the key allows us to translate a property of the map (the distance between the two dots being 18cm) into a property of the world (the distance between Rome and London being 1800km). The keys used in scientific models are often more complicated than the scale of a map and involve idealisations, approximations and analogies.

We are now in position to give an analysis of t-representation, i.e. of how models represent targets. Consider an Agent *A*. The agent chooses an object as the base of the representation and turns it into a *Z*-representation by adopting an interpretation *I*. The model *M* is the package of the object together with the interpretation *I* that turns it into a *Z*-representation. Model *M* then t-represents target *T* iff (i) *M* denotes *T* (and, possibly, parts of *M* denote parts of *T*); (ii) *M* is a *Z*-representation exemplifying certain *Z*-properties; (iii) *M* comes with a key *K* specifying how the *Z*-properties of the model are translated into another set of properties, and (iv) *M* imputes at least one of these other properties to *T*. This is the DEKI account of representation (where the acronym highlights the account's key features: denotation, exemplification, keying-up and imputation).

We now see how the first and the second part of this paper converge. The DEKI account requires a model to be an object that instantiates properties and can be interpreted. This is straightforward in the case of material models like wood models

of ships or ball-and-stick models of molecules. But most scientific models are not like this. Fibonacci's immortal rabbits are not material objects. But what object are they? The fiction view developed in Section 2 offers an answer that explains how claims about models can be true or false and how scientists can find out about these claims.
At this point the DEKI and the fiction view converge. DEKI explains how the features of model figure in scientific representation, and the fiction view – fleshed out in terms of games of make-believe, furnishes a notion of fiction that explains how models can be said to have properties that provide the input to the DEKI machinery.

4. Denotation

The above account appeals to denotation but remains silent about the nature of denotation. Denotation is a dyadic relation that obtains between certain symbols and certain objects. Proper names are paradigmatic examples of denoting symbols. For example, we can use the name "Socrates" to denote a particular individual, namely Socrates. Following Goodman and Elgin we assume that the domain of denotation includes more than just linguistic symbols. Other symbols like pictures, graphs, charts, diagrams, maps, and drawings represent their subjects by denoting them (Elgin 2010, 2). On the DEKI account also scientific models can denote.

Denotation is of course a time-honoured topic that raises many interesting questions. However, the nature of denotation is particularly pressing in the context of the fiction view of models. Discussions in the philosophy of language have predominantly been concerned with what we call the *relation problem*: in virtue of what does the denotation relation between a symbol and its object hold? In these discussions relatively little attention has been paid to the nature of the denoting symbols themselves, which by and large have been take for granted.[7]

[7] An important discussion is Kaplan's (1990) investigation into the nature of words. A more extensive investigation is found in Wetzel's (2009).

While this is not an unreasonable attitude in context of a discussion of language, the nature of the symbols is not something philosophers of science can be nonchalant about. Models denote their targets, meaning that models take the place of names and other denoting expressions in linguistic representation. But since models are said to be fictions, the view is committed to the claim that fictions denote target systems.

How can that be? How does, say, Fibonacci's model (which consists of fictional immortal rabbits) denote a real population of animals? As previously noted, denotation is a dyadic relation and only existent objects can enter into relations. If model systems are fictions, how could they possibly enter into such a relation and denote a target? The view that models denote forces the philosopher of science to deal with what we call the *identification problem*: what objects are models? While philosophers of language can assume that the nature of the symbols themselves is clear enough for their purposes, philosophers of science have a serious problem at their hands.

How one responds to this challenge depends on one's metaphysics of fiction and on how that metaphysics is integrated into an analysis of models. Realists about fictional entities assume that fictional entities exist, and so they could reply that denotation is not a problem. Whether realism can deliver a workable account of denotation is an interesting question, and the matter is not straightforward. For want of space we set this question aside and focus on our preferred option, fictional antirealism. On this view the characters and places described in fictions don't exist. But if the objects of fiction don't exist, they can't enter into relations with objects in the world and therefore can't denote. A theory that combines anti-realism about models with the view that models denote must therefore appear incoherent.[8] Antirealists can respond in at least two ways to this challenge.

The first response is to retract the idea that to represent a target system a model must really denote the target and hold that when scientists talk about a model denoting they actually talk *in pretence*. The claim that models denote can then be seen as part of the

[8] In passing we note that a realist theory that sees models as nonconcrete objects faces the same objection.

pretence of the model, which does not require that there be any genuine denotation between model and target. On this interpretation model denotation is pretend denotation, and scientists' claims about models denoting their targets must be reinterpreted as moves in a game of make-believe. Thus understood, claims about denotation can still have genuine truth conditions, but only when they are prefixed with a fictional operator.

There is a question whether pretend denotation is enough to account for scientific representation. Those who are fully invested into the pretence framework may well think that it is. Others may think that it is too feeble a notion to account for how science represents its objects and nothing short of "real" denotation between models and their targets is good enough. At this point we want to keep an open mind about this question and so it is desirable to also have a version of the fiction view that makes room for real denotation.

The second response dismisses pretend denotation and submits that models must genuinely denote. To come to an antirealist version of the fiction view that is consistent with genuine denotation, we must reconceptualise what a model is. This must be done in a way that turns the model into something that exists but without giving up the core-idea of anti-realism, namely that there are no fictional objects. One way of doing this is to associate the model with the content of the fiction *and* the text that describes the content rather than *only* with the fictional object that is described in the text. A model then is a tuple $M = [D, C]$, where D is the description of the model and C is the content of the description. It is important that C is the full content of the description, which consists both of a set of primary fictional truths and the implied fictional truths (i.e. the set of propositions that are specified by D together with the principles of generation). This ensures that models have content beyond what is stated in the model description, which is necessary for models to be vehicles of research.

C now takes the place of what one intuitively would call the 'model system' (such as Fibonacci's immortal rabbits). Because model-descriptions and their contents exist, models thus construed are *bona fide* objects (akin to fictional stories) that can enter into relations. On this view, the Fibonacci model is the description offered in the opening section of this paper and that description's content. The unit of the description along with its content exits and it therefore can denote, say, the rabbit population in the London Zoo. Some caution is needed at this point: model descriptions *themselves* don't denote anything because there is nothing that satisfies the properties and relations specified by *D* (there are no immortal rabbits!). It is *M*, the tuple [*D*, *C*], which can be the symbol that denotes the target.

Saying that Fibonacci's model denotes the rabbit population in the London Zoo requires that we have a solution to the relation problem. To pave the ground for discussion of this problem we first have to get clear on the nature of the problem. To that end we emphasise that models have no "in-built" denotation. Anything can, in principle, denote anything else, and models are no exception.[9] The denotation relation between model and target, if any, is extrinsic to the model: the model doesn't denote something in virtue of some special intrinsic features it has. A model is an object of the sort we discussed in the last section, and if it denotes it does so in virtue of entering into a relation with the target that is extrinsic to the model itself. In this respect models are no different from words, which have to be connected to objects in special ways for it to be the case that they denote (as we will see shortly), and they can, in principle, change their denotata when moving from one context to another.

Where does the denotation relation between models and their targets originate? We doubt that there is a fully general answer to this question. However, a look at scientific practice suggests that in many cases the denotation of a model piggy-backs on the denotation of denoting linguistic symbols. In our example, Fibonacci's model denotes what it does because we use the denoting expression "the rabbit population in the London Zoo". And the same happens in many other cases, where the representational function of models is navigated through language and models ultimately "borrow" their denotation from language.[10] Hence, a model denotes a target *T* in virtue of whatever it is that allows users of a language to use the term "*T*"

[9] In the context of representation, this point has been emphasized, among others, by Callender and Cohen (2006) and (Swoyer 1991).

[10] The expressions that give the model denotation need not be part of the model description. We noted that denotation is extrinsic, and so models need not (and often don't) denote whatever the model description contains. Fibonacci could have introduced his population model with immortal minotaurs, which would not make the model denote a minotaur population.

to denote *T*. The denotation of models is thereby reduced to that of the linguistic expressions referring to targets.

This observation has an important methodological implication: studying the denotation of models amounts to studying the referential practices of the scientific languages that are used in connection with models. This is a momentous task that is beyond the scope of this paper. But there is also no need to get into this here because the denotation of terms is a widely studied topic, and readers who wish to further pursue the matter can turn to the vast literature on the subject.[11]

5. Opposition

The fiction view of models has been criticised on grounds that have little to do with the specifics of the account. Some commentators think merely mentioning 'fiction' in a scientific context is an anathema and so the entire account got started on the wrong foot. Giere, for instance, accuses the fiction view of playing into the hands of science sceptics and irrationalists (2009, p. 257). Creationists and other anti-science activists may find great comfort, if not powerful rhetorical ammunition, in the fact that philosophers of science say that scientists produce fiction. This, so the argument goes, will be seen as a justification of the view that religious dogma is on par scientific knowledge. The fiction view of models thus undermines the authority of science.

We would be profoundly chagrined if our account of models would be used to muster support for intellectual aberrations like creationism. There is no absolute safeguard against misinterpretation, and the best intentions can be turned on their head. Darwinism has been used to justify eugenics, and millions have been killed in the name of a benevolent and loving God. The fear of misinterpretation can't count as an argument against the fiction view specifically. We hope that it has become clear from the exposition of the view, that, far from embodying irrationalist doctrines, it actually aims to explain what makes scientific modelling tick. The conclusion can therefore not be that we have to abandon the fiction view; the point to be taken away from

[11] A general introduction can be found in Lycan's (2008). For overview over the semantics of theoretical terms see Percival's (2000).

Giere's criticisms is that some care may be needed when dealing with the press office. As long as the fiction view of models is discussed in informed circles, and when popularised is presented carefully and with the necessary qualifications, it is no more dangerous than other ideas, which, when taken out of context, can be put to uses that would send shivers down the spines of their progenitors.

Acknowledgments

We would like to thank Bradley Armour-Garb and Fred Kroon for inviting us to participate in this project and for helpful comments on earlier drafts of this chapter.

References

Barberousse, Anouk and Pascal Ludwig. 2009. "Models as Fictions." In Fictions in Science: Philosophical Essays in Modeling and Idealizations, ed. Mauricio Suárez, 56-73. London: Routledge.

Berto, Francesco. 2011. "Modal Meinongianism and Fiction: The Best of Three Worlds." Philosophical Studies 152:

Callender, Craig and Jonathan Cohen. 2006. "There Is No Special Problem About Scientific Representation." Theoria 55:7-25.

Contessa, Gabriele. 2010. "Scientific Models and Fictional Objects." Synthese 172:215-229.

Elgin, Catherine Z. 1983. With Reference to Reference. Indianapolis:

Elgin, Catherine Z. 2010. "Telling Instances." In Beyond Mimesis and Convention. Representation in Art and Science, ed. Roman Frigg and Matthew C. Hunter, 1-17. Berlin and New York: Springer.

Everett, Anthony. 2013. The Nonexistent. Oxford: Oxford University Press.

Frigg, Roman. 2010. "Models and Fiction." Synthese 172 251-268.

Frigg, Roman and James Nguyen. 2016. "The Fiction View of Models Reloaded." The Monist 99:225–242.

Frigg, Roman and James Nguyen. 2017. "Models and Representation." In Springer Handbook of Model-Based Science, ed. Lorenzo Magnani and Tommaso Bertolotti, Dordrecht, Heidelberg, London and New York: Springer.

Frigg, Roman and James Nguyen. 2018. "The Turn of the Valve: Representing with Material Models." European Journal for Philosophy of Science 8:205-224.

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

Giere, Ronald N. 2009. "Why Scientific Models Should Not Be Regarded as Works of Fiction." In Fictions in Science. Philosophical Essays on Modelling and Idealization, ed. Mauricio Suárez, 248-258. London: Routledge.

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: Hacket.
Kaplan, David. 1990. "Words." Proceedings of the Aristotelian Society, Supplementary Volumes 1:93-119.
Kroon, Frederick. 2011. "The Fiction of Creationism." In Truth in Fiction, ed. Franck Lihoreau, 153-172. Heusenstamm: Ontos.
Levy, Arnon. 2015. "Modeling without Models." Philosophical Studies 152:781–798.
Lopes, Dominic. 2004. Understanding Pictures. Oxford: Oxford University Press.
Lycan, William G. 2008. Philosophy of Language. A Contemporary Introduction. New York and London: Routledge.
Parsons, Terrence. 1980. Nonexistent Objects. New Haven: Yale University Press.
Percival, Philip. 2000. "Theoretical Terms: Meaning and Reference." In A Companion to the Philosophy of Science, ed. William H. Newton-Smith, 495-514. Oxford: Wiley-Blackwell.
Salis, Fiora. 2016. "The Nature of Model-World Comparisons." The Monist 99:243-259.
Salis, Fiora and Roman Frigg. forthcoming. "Capturing the Scientific Imagination." In The Scientific Imagination. Philosophical and Psychological Perspectives, ed. Peter Godfrey-Smith and Arnon Levy, Oxford: Oxford University Press.
Swoyer, Chris. 1991. "Structural Representation and Surrogative Reasoning." Synthese 87:449-508.
Thomasson, Amie L. 1999. Fiction and Metaphysics. New York: Cambridge University Press.
Thomasson, Amie L. forthcoming. "If Models Were Fictions, Then What Would They Be?" In The Scientific Imagination. Philosophical and Psychological Perspectives, ed. Peter Godfrey-Smith and Arnon Levy, Oxford Oxford University Press.
Thomson-Jones, Martin. 2010. "Missing Systems and Face Value Practise." Synthese 172:283 - 299.
Thomson-Jones, Martin. forthcoming. "Realism About Missing Systems." In The Scientific Imagination. Philosophical and Psychological Perspectives, ed. Peter Godfrey-Smith and Arnon Levy, Oxford Oxford University Press.
Toon, Adam. 2012. Models as Make-Believe. Imagination, Fiction and Scientific Representation. Basingstoke: Palgrave Macmillan.
Walton, Kendal L. 1990. Mimesis as Make-Believe: On the Foundations of the Representational Arts. Cambridge/MA: Harvard University Press.
Weisberg, Michael. 2013. Simulation and Similarity: Using Models to Understand the World. Oxford: Oxford University Press.
Wentzel, Linda. 2009. Types and Tokens: On Abstract Objects. Cambridge/MA: MIT Press.
Zalta, Edward N. 1988. Intensional Logic and the Metaphysics of Intentionality. Cambridge/MA: MIT Press.