

The Nature of Model-World Comparisons

Fiora Salis

This is the penultimate version of a paper that will appear in a special issue of *The Monist* on Scientific Fiction Making. The final version will be available at <http://monist.oxfordjournals.org>; citations should be to the final version.

Abstract: Upholders of fictionalism about scientific models have not yet successfully explained how scientists can learn about the real world by making comparisons between models and the real phenomena they stand for. In this paper I develop an account of model-world comparisons in terms of what I take to be the best antirealist analyses of comparative claims that emerges from the current debate on fiction.

1. Introduction

In recent years philosophers of science have developed different versions of a fruitful analogy between theoretical models and the fictions of literature and the arts. One special problem for fictionalism about theoretical models consists in making sense of how scientists can learn about the real world by making comparisons between models and real systems. This is the problem that I will address in this paper.

Typically, when scientists study a particular part or aspect of the world that would be too complex to be analysed in full detail they construct a simplified version of it, i.e. a scientific model. Some models are concrete physical systems – or *material models* – such as Watson and Crick’s metal model of DNA and Phillips’s hydraulic model of economy. Others are non-actual systems – or *theoretical models* – that do not exist as physical concrete systems – such as Newton’s model of the solar system

and the Rutherford-Bohr model of the atom. Godfrey-Smith (2006) and Weisberg (2007, 2013) suggest that scientific modelling involves two distinct relations: a specification of a model system and a set of comparisons between the model system and the real world phenomena it stands for¹. A model system is chosen as the object of study because it is less complicated than its target and because by using the model system to represent its target we can learn about the latter. Thus, scientists construct, develop, and test models in order to discover truths about their targets. This prompts the question of how learning with a model is possible. Answering this question requires an account of model-world comparisons.

There are three different sorts of comparisons we can make. First, we can compare two real physical systems, e.g. the water in Phillips' hydraulic model of economy flowing from the treasury tank to other tanks with how the money in the real economy flows from the real treasury to other areas such as health care or education. Second, we can compare two non-actual systems, e.g. an electron's orbit around the nucleus in the Rutherford-Bohr model of the atom with the motion of a planet around the sun in Newton's model of the solar system. Third, we can compare real systems with non-actual systems, e.g. the growth of a population of rabbits in a real ecosystem with the way in which a non-actual population in a model of population growth behaves. Comparisons of the first sort are unproblematic to the extent that they involve ordinary physical systems and properties. But how can we compare non-actual systems with each other, and non-actual systems with actual concrete systems?

Philosophers of science have worried about model-world comparisons, but they have not yet successfully accounted for them. For example, Giere (1988)

¹ Giere (1988) originally formulated this distinction within a broader account of scientific theories.

proposes to think about non-actual model systems as abstract objects standing in a relation of similarity in certain respects and to certain degrees with concrete systems. Hughes (1997), however, notices that there is no obvious sense in which abstract objects can be similar to concrete objects. And Thomson-Jones (2010) argues that if similarity is interpreted as a matter of sharing properties, abstract objects cannot have the sort of spatio-temporal properties that real systems have. Weisberg (2013) analyses model-world comparisons in terms of a higher order similarity relation between interpreted mathematical representations of models and targets. This higher order relation, however, supervenes on a lower order similarity relation between non-actual systems and concrete targets that, as I will argue later in this paper, still needs to be fully explained.

Godfrey-Smith (2009) suggests that the analogy between models and fiction may deliver a solution to the problem of comparative claims. However, he does not develop a full-blown account of model-world comparisons. Frigg (2010) suggests that comparative claims come down to comparing properties of non-actual systems and properties of real systems. Godfrey-Smith (*ibid.*) argues that this is problematic because the properties of a non-actual system are un-instantiated. Toon (2012) and Levy (2015) argue that modelling involves imaginative descriptions of real world targets and deny that there are any model-world comparisons because there are no model systems. Yet, scientists do engage in model-world comparisons. So, they face a different version of the problem, one of comparing the content of an imaginative description of a real thing with the thing itself, e.g. the growth of a population of rabbits in a real ecosystem with the way in which the same population behaves in a model of population growth. As I will argue later, I don't think this makes the problem easier and it still needs a solution.

In this paper I aim at offering a plausible account of model-world comparisons that naturally follows from Walton's (1990) theory of fiction. I will present this theory in Section 2. I will assume what I take to be the best account of comparative claims that follows from this theory coherently with fictional antirealism in Section 3. In light of this discussion, I will critically assess the current accounts from the debate on modelling in Section 4. Finally, in Section 5 I will advance an alternative solution in terms of the best antirealist analyses produced within the debate on fiction.

2. Walton's aesthetic notion of fiction

The key to a rigorous construal of the analogy between models and fiction is in the aesthetic notion of fiction that applies to works of imaginative narration such as Shakespeare's *Othello* and Tolstoy's *Anna Karenina*. Literary fictions often contain statements that, when judged from a real world perspective, are evaluated as false and that are often about objects that do not exist as concrete spatiotemporal entities. For example, in Orwell's *1984* London is the capital of a fascist state, and in Doyle's *A Study in Scarlet* Sherlock Holmes lives at 221B Baker Street. The first statement is false because the name 'London' refers to the capital of the UK but this does not satisfy the predicate 'being the capital a fascist state'. The second statement is false – or untrue – because the name 'Sherlock Holmes' does not refer to any real individual and 221B Baker Street does not exist. The aesthetic notion, however, is not defined in terms of falsity and non-existence. An historical account does not become a literary work of fiction because it contains some false claims or apparent reference to non-existent individuals. And a fictional work does not become a work of non-fiction because it contains some true claims and reference to real individuals.

Originally Walton spelled out an explanatory connection between make-believe and fiction, where by ‘make-believe’ he means ‘the use of (external) props in imaginative activities’ (1990, 67). According to Walton, works of fiction function as props in games of make-believe, just like dolls, toy trucks and teddy bears function as props in children’s games. Props are real objects that generate fictional truths in virtue of there being a prescription to imagine something, i.e. a social convention either explicitly stipulated or implicitly understood as being in force within a certain game.

Fictional truth – or fictionality – is a property of those propositions that are among the prescriptions to imagine of a certain game. Thus, imagined propositions have correctness conditions whereby only some of them are licensed by an author’s prescriptions to imagine. The proposition that Holmes lives at 221B Baker Street is fictionally true because of Doyle’s prescriptions to imagine. The proposition that he is a cleverly disguised dog, however, is fictionally false because this is not among the imaginings prescribed by the story.

Furthermore, fictional truths divide into primary truths and implied truths, where the first are generated directly from the text (the prop), and the second are generated indirectly from the primary truths via general principles that are taken for granted and standard rules of inference that can be reality-oriented. For example, from the primary fictional truth that Sherlock Holmes lives in Baker Street and our knowledge of London’s geography we can infer the implied fictional truth that Holmes lives nearer to Paddington Station than to Waterloo Station.

From this aesthetic notion flows a natural interpretation of fictional discourse – our discourse about fictional objects, events and situations – which is crucial to the analysis of comparative claims that I will develop in Section 3. Consider an utterance of

1. Holmes lives at 221B Baker Street.

This can be interpreted in two ways corresponding to two kinds of fictional discourse.

First, there is *internal* or *intra-fictional discourse* that we perform from *within* the imagined context of the fiction, from a participatory or internal perspective. In this case, (1) is interpreted as a natural continuation of our imaginative engagement with the story and the propositional attitude we have towards (1) is imagination. *Prima facie*, an utterance of (1) is an assertion that can be evaluated for truth and falsity. We know, however, that Holmes does not exist and that he does not live in Baker Street. We merely imagine that he does. Thus, an utterance of (1) is better interpreted as an assertion (or pseudo-assertion) that is only fictionally true.²

Second, there is *external* or *meta-fictional discourse* that we perform from *without* the imagined context of the fiction, from a descriptive or external perspective. In this case we exit the fiction and assume an attitude of belief rather than imagination. We do not believe that Holmes lives in Baker Street, but we do believe that according to the fiction he does. In this case, an utterance of (1) is interpreted as an assertion that

² Walton suggests that when we engage in internal discourse involving referring terms without referents we pretend to express a proposition without really expressing anything. This semantic interpretation has been widely criticized for reasons that go beyond the scope of the present discussion (cf. Richard 2000, and Stanley 2001). On a more plausible interpretation, an utterance of (1) expresses a proposition that, depending on your semantic preferences, can be a general proposition (e.g., Currie 1988) or a gappy proposition of the form ‘*x* lives at 221B Baker Street’ (e.g., Braun 2005).

can be evaluated for genuine truth with respect to the real world. Doing this requires prefixing the original sentence with an ‘according to the fiction’ operator:

2. According to *A Study in Scarlet*, Holmes lives at 221B Baker Street.

An utterance of (2) is true if and only if the proposition embedded within the ‘according to the fiction’ operator is among the imaginings prescribed by *A Study in Scarlet*, and false otherwise.³

3. Comparative claims

Now we can analyse comparative claims involving fictional characters and real individuals. The considerations I will make in this section will work as a blueprint for the critical discussion of existent proposals from the literature on models in Section 4

³ There are two main interpretations of utterances of implicitly prefixed sentences such as (2). First, they can be interpreted as offering a *replacement* of the original sentence. After uttering (1) we can be pressed to clarify what we really mean. If we intend to convey something that is really true we can replace it with (2). Second, they can be interpreted as offering a *paraphrase* of the original sentence, which is construed as being equivalent in truth conditions to the new prefixed sentence. The paraphrase is supposed to reveal the logical form of the original sentence and its ontological commitments. This is a theoretically highly loaded interpretation that raises several controversial issues (cf. Yablo 1999). The replacement strategy, however, is theoretically less demanding (and therefore preferable) because it is not supposed to agree in truth conditions with the original sentence and it does not commit to the problematic notion of logical form.

and for the development of a new proposal in Section 5. Consider:

3. Zaphod Beeblebrox is more narcissistic than Morris Zapp.
4. Putin is more neurotic than Raskolnikov.

There is not one fictional story according to which Zaphod Beeblebrox (from Douglas Adams' *The Hitchhiker's Guide to the Galaxy*) is more narcissistic than Morris Zapp (from David Lodge's *Changing Places*), or one fictional story according to which Putin is more neurotic than Raskolnikov (from Dostoyevsky's *Crime and Punishment*). Yet, on the internal reading, (3) and (4) can be naturally interpreted as performed within an extended imagined context resulting from the combination of the original prescriptions to imagine. In Walton's own words, they 'can be thought of as contributions to unofficial games which combine in natural ways games authorized for the various works' (*ibid.*, 407). By 'unofficial games' he means games that are not originally licensed by an author's prescriptions to imagine, but that are generated for special purposes in special circumstances (*ibid.*, 405-ff.).

What can we say about utterances of (3) and (4) as performed from an external perspective? Upholders of *fictional realism* postulate different sorts of fictional entities as the bearers of the names apparently referring to them and offer different analyses. Neo-Meinongians think that they are concrete non-existent objects (Parsons 1980), or possible but non-actual objects (Berto 2011). Abstract object theorists claim that they can be construed as abstract eternal Platonic entities (Zalta 1983, 1988), or as abstract artefacts akin to other social constructs (Thomasson 1999). I will not discuss Neo-Meinongianism because there is no correspondent theory of models.

However, there are abstract object theories of models sharing many features with abstract object theories of fiction. So, I will focus on these theories here.⁴

Abstract object theorists interpret utterances of (3) and (4) as having the relational form Rab , where a and b are two objects standing in relation R . This requires postulating abstract entities as the referents of the names that apparently refer to them. However, fictional characters are supposed to have properties that abstract objects cannot have. For example, Zaphod and Morris are supposed to be narcissistic, and Raskolnikov is supposed to be neurotic. Abstract object theorists advance two main strategies to account for this phenomenon.

The first strategy appeals to Thomasson's (1999) distinction between two different kinds of sentential contexts: a fictional context and a real context. Abstract objects can have properties such as being a fictional character or having been created by Douglas Adams. But they cannot really have properties such as being human, being neurotic or being narcissistic. At most, they can have these properties in a fictional context. In Walton's terms, we can only imagine them as having these properties from within the fiction, from an internal perspective.

The second strategy appeals to a distinction between two kinds of predication or two kinds of relations between properties and individuals. Van Inwagen (1977) suggests that our discourse about fictional characters is ambiguous between genuine predication and the mere ascription of properties to abstract entities. We genuinely predicate the property of being neurotic to Putin (a concrete individual), but we merely ascribe the property of being neurotic to Raskolnikov and the property of being narcissistic to Zaphod and Morris. Similarly, Zalta (1983) submits that fictional

⁴ For a critical overview of the controversies surrounding the ontology of fictional entities see Salis (2013).

objects *encode* the sort of properties that only concrete objects can have, while ordinary objects *exemplify* them. Putin exemplifies the property of being neurotic, Raskolnikov merely encodes the same property and Zaphod and Morris encode the property of being narcissistic.

A natural way to understand the notions of ascription and encoding is as varieties of representation. Abstract objects are represented, in the imagination, as having properties that they do not really have⁵. Comparing abstract objects and concrete objects with respect to the sort of properties that only concrete objects can have requires that we imagine the first as having properties that they do not really have. Thus, whether or not we accept fictional entities into our ontology, the key to a genuine solution seems to reside in our imaginative engagement with fictions. Coherently with this idea, upholders of fictional antirealism reject fictional entities and analyse comparative claims in two different ways corresponding to what I call *Analysis₁* and *Analysis₂*.

The key to *Analysis₁* is in the notion of an extended fiction or, in Walton's terminology, an unofficial game of make-believe. Let us begin by considering (3). There is not one story according to which Zaphod is more narcissistic than Morris. In order to offer genuine truth conditions for this statement we must allow fictional operators to *agglomerate*:

⁵ In fact, Van Inwagen proposes to think about ascription as a three-place relation between an abstract object, a property and a place, where the latter is 'either a work of fiction (such as a novel, short story, or narrative poem) or a part or a section thereof, even a part or a section that is so short as to be conterminous with a single (occurrence of a) sentence or clause.' (1977, 305)

5. According to *The Hitchhiker's Guide to the Galaxy* and *Changing Places*, Zaphod is more narcissistic than Morris.

The combination of the two stories generates an unofficial game of make-believe wherein the fictional truth that Zaphod is more narcissistic than Morris is implied by the primary truths of the combined stories. In the combined stories Zaphod and Morris exist and certain relevant information about the narcissism of Zaphod and the narcissism of Morris further implicate that Zaphod is more narcissistic than Morris.

What happens if there is a conflict between the combined stories? Say that according to *The Hitchhiker's Guide to the Galaxy* there is a president of the universe while according to *Changing Places* there is not. This problem does not arise in the fragmentary unofficial game involving only Zaphod's and Morris' narcissism. In this case, we play an unofficial game combining only the primary truths regarding Zaphod's and Morris' narcissism. If the game is further pursued and conflicts between the two stories arise, they can be confronted in different ways in different contexts.⁶

Now let us consider (4). *Crime and Punishment* does not say or imply anything about Putin and his being neurotic. On the internal construal we engage in an imaginative activity involving Putin and Raskolnikov. In this case, we evaluate the utterance with respect to an extended fiction, or unofficial game, that includes the relevant fragments of information about Raskolnikov and about Putin which is imported into the game from the novel and from the real world respectively. Hence, the external reading of (4) requires an extended fictional operator:

⁶ See Walton (1990, 408) for similar considerations on a different case involving inter-textual identity.

6. According to the extended fiction, Putin is more neurotic than Raskolnikov.

An utterance of (6) is genuinely true, not just fictionally true, on the assumption that there is an extended fiction, or unofficial game of make-believe, according to which Putin is more neurotic than Raskolnikov. This fragmentary game only involves information about Putin and Raskolnikov. Hence, there is no conflict between *Crime and Punishment* and the real world. However, if the the game is further pursued conflicts may arise that will need to be solved according to further principles of generation to be established in different ways in different contexts.

Analysis₂ appeals to a genuine comparison between *degrees* of properties, which are mathematical entities that can be individuated on a scale of measurement (cf. Walton 1990, 413-414). Let us consider (3) again. When we say that Zaphod is more narcissistic than Morris we say (or imply) that there are certain degrees of narcissism i and j such that $i > j$, Zaphod has narcissism to degree i , and Morris has narcissism to degree j . This analysis is problematic because Zaphod and Morris do not exist, so they cannot really have degrees of narcissism i and j . In other words, Zaphod and Morris do not instantiate the relevant degrees of properties. And (3) is still false.

Appealing to external quantification over degrees of properties embedding internal quantification over stories can solve the problem. The improved analysis becomes:

7. There are some degrees of narcissism, i, j , such that $i > j$, according to *The Hitchhiker's Guide to the Galaxy* Zaphod has narcissism to degree i , and according to *Changing Places* Morris has narcissism to degree j .

(7) is true on the assumption that there are degrees of narcissism i, j standing in the greater-than relation, and that according to *The Hitchhiker's Guide to the Galaxy* Zaphod has narcissism to degree i , and according to *Changing Places* Morris has narcissism to degree j .

The same analysis applies to (4). In this case, we only need to appeal to external quantification over degrees of neuroticism embedding internal quantification over one fictional story:

8. There are some degrees of neuroticism, g and h , such that $g > h$, Putin has neuroticism to degree g , and according to *Crime and Punishment* Raskolnikov has neuroticism to degree h .

Some philosophers of fiction prefer Analysis₁ to Analysis₂. Paradigmatically, Sainsbury (2010) makes two different critical considerations. First, he claims that it is doubtful that authors of fiction make genuine *de re* statements about mathematical entities (degrees of properties) that relate to fictional objects.⁷ However, we do not need to claim that they do. Analysis₂ only requires that they say things that imply such statements. Second, he worries that ‘the proposed approach will not generalize’ (*ibid.*, 124). The case he has in mind is a variation of Kafka’s *The Metamorphosis* where Gregor Samsa is turned into a many-legged creature with an unspecified number of legs. In this case Gregor has more legs than Sainsbury has, but we do not

⁷ By ‘*de re* statements’ philosophers of language mean different things, but in this case Sainsbury simply means statements involving reference to specific (mathematical) objects.

know how many more: ‘there is no number $i > 2$, such that according to (modified) Kafka, Gregor has i legs whereas I have just 2’ (*ibid.*, 124). However, comparing numbers does not always require that we compare definite numbers. We often see arguments such as ‘there exists a number y such that $y > x$ for some specific number x ’. No one requires that there is a definite value of y for the argument to go through.

So, in this section I have done three things. First, I have paved the way for an understanding of comparative claims that will be crucial for a critical assessment of the existent accounts from the literature on models and for the positive development of an alternative view. Second, I did this by avoiding realism not because it is an implausible view but because it does not deliver a solution to the problem of comparative claims that requires fictional entities. Third, I showed that the alternative antirealist analyses have the resources to offer an account that dispense with fictional entities. They recognize as fundamental our imaginative engagement with fictions, yet they also allow us to assume an external perspective for which believe, rather than imagination, is the appropriate attitude. Furthermore, by embedding the problematic statements in the fictional operator they can also explain how they can be genuinely true without committing us to an ontology of fictional entities.

4. Three views of model systems and model-world comparisons

In this section I will consider how the above discussion can contribute to a better understanding of model-world comparisons. There are three main different views on the nature of model systems. Upholders of the *abstract view* claim that they are abstract entities. Upholders of *indirect fictionalism* suggest that model systems are akin to fictional characters. Upholders of *direct fictionalism* reject model systems and

claim that model descriptions are directly about the world. In this section I will critically assess these views in relation to the problem of comparative claims.

Let us begin with the abstract view, which was originally advanced by Giere (1988) within a broader account of scientific theories. Giere notices that model descriptions used in mechanics textbooks are not satisfied by any real systems. For example, there are no frictionless planes, no ideal pendulums, and no bodies subject only to a central gravitational force. He proposes to regard model systems as ‘abstract entities having all and only the properties ascribed to them in the standard texts’ (*ibid.*, 78). More specifically, he regards them as ‘*socially* constructed entities’ having ‘no reality beyond that given to them by the community of scientists’ (*ibid.*). Hence, model systems are ontologically on a par with the abstract artefacts of Thomasson’s (1999) artefactual theory of fiction⁸.

Giere suggests that model-world comparisons are theoretical hypotheses about the similarity relation between a model system and a target, which is implicitly specified according to certain ‘*respects* and *degrees*’ (*ibid.*, 81). Consider his example:

9. The positions and velocities of the earth-moon system are very similar to those of a two-particle Newtonian model with an inverse square force.

⁸ Giere (2009, 2004, 2010) recognizes that models and fictions are ontologically on a par. Yet, he rejects the analogy between models and fiction mainly on the basis of considerations about their seemingly distinct functions. See Frigg (2010b, 280-81) for responses.

A hypothesis is a proposition that is a candidate for truth. Yet, as Hughes (1997) and Thomson-Jones (2010) point out, abstract objects cannot have the sort of properties that ordinary objects have. That is, the abstract two-particle Newtonian system cannot have any positions or velocities. Thus, (9) cannot be true. As I pointed out in the previous section, there is no solution to this problem in Thomasson's version of fictional realism. As far as Giere's realism is on a par with Thomasson's view, his view will confront the same problems.

Weisberg (2013) develops a version of the abstract view according to which models are composed of structures and construals. Structures can be concrete (for material models), mathematical or computational (for theoretical models). Construals involve an assignment, a modeller's intended scope, and two kinds of fidelity criteria, i.e. dynamical and representational. For example, Weisberg identifies the Lotka-Volterra population model of predation with the following interpreted differential equations:

$$10. \quad dV/dt = rV - (\alpha V)P$$

$$11. \quad dP/dt = b(\alpha V)P - mP$$

The interpreted equations represent the growth rates of two hypothetical populations. Their construal involves the following assignments: let V stand for the size of some prey population, let P stand for the size of some predator population, let t stand for time, and let r , α , b , and m stand for real parameters describing the interaction of the two populations. The scope of the model is limited to selected elements including the size of the two populations, their birth rates and death rates, the prey capture rate, and the number of prey captured to produce the birth of a predator. Fidelity criteria

describe how similar the model must be to the target. For example, the population size of the hypothetical preys and predators must approximate a value that is $\pm 10\%$ of the value of the population size of real preys and predators.

Weisberg analyses model-world comparisons in terms of an isomorphism between theoretical models and mathematical representations of targets, which partly supervenes on model construals to the extent that ‘When context or scientific goals change, the construal will change, and aspects of the relation will change.’ (*ibid.*, 149). On his weighted feature-matching account of similarity this is represented as the intersection of features of the model and features of the target. Features divide between attributes, which are properties and patterns, and the underlying causal mechanisms that generate them. So, for example, the Lotka-Volterra model involves attributes such as equilibrium abundance and maximum population size and mechanisms such as the interaction of the predator and prey populations. Weisberg indicates the following equation as the core of his account:

$$S(m, t) = \frac{|M_a \cap T_a| + |M_m \cap T_m|}{|M_a \cap T_a| + |M_m \cap T_m| + |M_a - T_a| + |M_m - T_m| + |T_a - M_a| + |T_m - M_m|}$$

where ‘S’ stands for the similarity relation between model m and target t , ‘ M_a ’ and ‘ M_m ’ stand for the attributes and mechanisms of the model, and ‘ T_a ’ and ‘ T_m ’ stand for the attributes and mechanisms of the target. The equation represents the way in which Weisberg thinks of similarity as ‘the ratio of features shared to those not shared’ (*ibid.*, 148), with shared ones in the nominator and non-shared ones in the denominator. So, when model and target share many features the value of S

approaches 1, when they are identical $S = 1$, and when they do not share any feature $S = 0$.

Weisberg claims that fictionalism about models is a mere folk ontology that should ultimately be rejected when assessing model-world comparisons. Yet, presumably, the assessment of model-world comparisons between interpreted mathematical representations of models and targets will depend upon whether the hypothetical objects involved in the former can really have the sort of features they are supposed to share with the real world objects involved in the latter. For example, assessing the similarity relation between the Lotka-Volterra model and the interpreted mathematical representation of specific target populations will depend upon whether the hypothetical populations apparently referred to in the model construal can really have the features they are supposed to share with the real populations referred to in the interpreted mathematical representation of the target.

On a realist interpretation, Weisberg's analysis effectively collapses into Giere's analysis, which does not deliver a solution. Model populations are abstract objects that cannot have the sort of attributes that real populations have and cannot undergo any causal mechanisms. In this case, the statement 'the population size of the hypothetical preys and predators is similar to the population size of real preys and predators' is false. On an antirealist interpretation, the hypothetical prey and predator populations of the Lotka-Volterra model do not really exist. So, they cannot really share any properties with real world populations. We can only imagine that they do. As we know from the previous section, fictional antirealists can account for the intuition that comparative claims involving apparent reference to fictional objects are genuinely true by taking as fundamental our imaginative engagement with fictions

and by embedding the problematic statements within the appropriate fictional operators. I will develop this idea in the next section.

For now, I would just like to notice that, *pace* Weisberg, fictionalism about models cannot be rejected when assessing model-world comparisons. Fictions are crucial to his understanding of the model world relation to the extent that model construals involve apparent reference to hypothetical systems having features (attributes and mechanisms) that only concrete objects can have. Realism does not deliver any solution to this problem. Antirealism, however, will deliver the right answer.

Now let us consider the indirect fictionalist view according to which model systems are akin to fictional characters. Godfrey-Smith originally suggested that the objects described by modellers ‘might be treated as similar to [...] the imagined objects of literary fiction’ (2006, 735) and that a natural way to describe model systems is as ‘creatures of the imagination’ (2009, 101). He states his preference for a naturalistic and deflationary approach to the ontology of models, but he does not offer a full-blown theory of the nature of model systems and of the model-world relationship.

Frigg (2010a) develops an indirect fictionalist view in terms of Walton’s (1990) theory of fiction. He thinks of model descriptions as props in a game of make-believe and interprets the modelling practice as involving a scientist’s development of a hypothetical system starting from the basic set of primary assumptions of the model (its primary truths) and certain general laws. For example, the implied truth that the earth moves on an elliptical orbit is generated from the primary truths of the Newtonian model and the laws of classical mechanics.

Frigg recognizes that comparative statements seem to be genuinely true and

that they seem to lead to genuine knowledge of the world. He thinks that he can avoid the controversial issues surrounding the ontology of fictional entities by claiming that when scientists engage in model-world comparisons they compare features of a model system with features of its real world target. For example, when we compare a real population of rabbits with a hypothetical population in the predator-play model we can say things like:

these populations possess certain relevant properties which are similar in relevant respects ... and the statement making this comparison is true iff the statement comparing the properties with each other is true.
(Frigg, 2010a, 263-264).

On his view, comparing features of a model system with features of a real system is unproblematic because

the problems that attach to [comparative statements in the context of scientific modelling] have nothing to do with issues surrounding fictional discourse. (*ibid.*, 263)

This should free us from the problem of how to compare something with a nonexistent object, which on Frigg's view 'does not seem to make sense' (*ibid.*, 263). Godfrey-Smith (2009) notices that the properties of the model system that are compared to those of the real world target system are not instantiated. There is no rabbit population instantiating the properties that are compared to those of a real rabbit population. So, on this interpretation, model-world comparisons are still false.

Finally, let us consider direct fictionalism. Toon (2012) originally advanced a direct fictionalist account of models in terms of Walton's (1990) distinction between two sorts of imaginings. Some imaginings are about real objects, e.g. H.G. Wells' *The War of the Worlds* prescribes imagining *of* St Paul's Cathedral that its dome is heavily damaged. Other imaginings are about fictional characters, e.g. Stoker's *Dracula* prescribes imaginings that Count Dracula (a fictional individual) is so and so. Model descriptions of models with targets prescribe imaginings of the first sort, e.g. the model description of the ideal pendulum prescribes imaginings about a particular real pendulum. Model descriptions of models without targets prescribe imaginings of the second sort – where models without targets include models of discredited entities such as phlogiston, models of synthetic molecules that have not yet been created in a laboratory, etc. Toon does not develop an account of model-world comparisons but direct fictionalism entails that model-world comparisons consist in comparing the content of an imaginative description of a real system with the system itself. There are four problems that I can see with this approach.

First, Toon submits that the direct approach has the advantage of avoiding a controversial ontology of fictional entities that other accounts incur by prescribing imaginings about real things. He appeals to Walton's theory of make-believe, but he does not seem to recognize that the theory naturally combines with fictional antirealism, which rejects fictional entities. One can assume the same approach in the case of models and thereby avoid any controversial ontological commitments. Once we do this, the purported advantage of the direct approach disappears.

Second, the direct approach does not really get rid of the ontological controversies surrounding fiction. On this view, models with targets prescribe imagining *of* some particular real system that it is different from the way it really is.

This is tantamount to say that model descriptions prescribe imagining a fictional situation – or scenario – wherein a real system has some features that it does not really have. For example, the model descriptions of the ideal pendulum prescribes imagining *of* a real pendulum that it is a point mass suspended from a massless string swinging in a uniform gravitational field and loosing no energy to friction or air resistance. So, the ontological problem is reinstated as a problem about the nature of the fictional scenarios specified by model descriptions.

Third, one may doubt the plausibility of the very idea that model descriptions of models with targets prescribe imaginings about particular targets. Toon recognizes that ‘there are clearly many models that represent no *particular* object or event’ (2012, 76). About Bohr’s model of the hydrogen atom he writes that ‘presumably, it does not represent any particular hydrogen atom (although it might be used to do so)’ and that ‘we might think that the model represents a type of object or event’ (*ibid.*). Indeed, it would be implausible to argue that there was a particular hydrogen atom that Bohr’s model description is about. But the same consideration can be made about the ideal pendulum. The model was developed for the purpose of learning about any system of the relevant type, including a bob bouncing on a string, a person moving on a swing, a wrecking ball on a crane’s cable, or a pendulum clock. Of course, models (not model descriptions) are used to model specific targets. But the model description of the ideal pendulum is not about any particular pendulum just like the model description of Bohr’s model is not about any particular hydrogen atom. More plausibly, model descriptions of models with targets prescribe imaginings about hypothetical systems.

Fourth, Toon does not develop an account of how we learn about the real world through the practice of modelling. He submits that learning about a theoretical model ‘is not a matter of learning facts about any object. Instead, it is a matter of

discovering what is fictional in the world of the model' (*ibid.*, 47). The ultimate goal of modelling, however, is to learn *facts* about the real world. In order to do this, scientists compare models with targets. Toon claims that in some cases similarity seems to play a role. He argues that

'principles of generation often link properties of models to properties of the system they represent ... If the model has a certain property then we are to imagine that system does too. If the model is accurate, then the model and system will be similar in this respect' (*ibid.*, 68-69).

Notice, however, that in this way Toon explains only how the model and the target can share certain properties *within* the make-believe. In this case the relevant comparisons are only fictionally true rather than genuinely true. And we do not know how to export knowledge acquired in the make-believe into knowledge of the real world.

Levy advances a similar version of direct fictionalism in terms of Walton's (1993) notion of *prop-oriented make-believe*, which consists in imagining of a specific concrete object (the prop) that it is different from the way it actually is. On Levy's view, '[m]odels are imaginative descriptions of real-world phenomena' (*ibid.*, 797), where the real world phenomena are props that generate fictional truths in virtue of a scientist's prepared descriptions. However, Levy develops a different account of how modellers learn about the real world in terms of Yablo's (2014) notion of partial truth. Roughly stated, a proposition is partly true if and only if it has parts that are wholly true. The propositions expressed by model descriptions are false, yet they are about a real target and they are partly true if some of the things they say about the

target are true.

One major problem for this view consists in explaining how we can clearly distinguish and quarantine those parts of model descriptions that are true from those that are false. Presumably, this is an empirical question to be answered within scientific practices. However, it is reasonable to suppose that such quarantining may not always be possible.⁹ Furthermore, Levy rejects the idea that learning with models involves engaging in model-world comparisons. Yet, scientists do make these comparisons and any philosophical account of model-based science should include an analysis of them.

5. Solving the problem

In the previous section I showed that upholders of the three main approaches to the nature of scientific models have not yet successfully accounted for model-world comparisons. Now it's time to develop a positive solution in terms of what I take to be the best antirealist analyses that emerge from the current debate on fiction and that I presented in Section 3. Let us begin by considering a comparison between fictional systems such as the hypothetical systems described in Newton's model of the solar system and in the Rutherford-Bohr model of the atom:

12. Electrons orbit around the nucleus just like planets orbit around the sun.

An utterance of (12) can be understood as produced from within or from without an

⁹ See Bokulich (2008, 2009, 2012, 2016) for several paradigmatic examples.

imagined context. Since there is not one model according to which electrons move around the nucleus just like planets move around the sun, the utterance can only be fictionally true from *within* an extended imagined context, i.e. from an internal or participatory perspective. In this case we naturally assume an attitude of imagination towards its content.

However, there is also a sense in which (12) seems to be true from *without* the extended imagined context, i.e. from a descriptive and non-participatory perspective. Analysis₁ can account for this intuition by prefixing (12) with fictional operators that agglomerate:

13. According to the Rutherford-Bohr model of the atom and to the Newtonian model of the solar system, electrons orbit around the nucleus just like planets orbit around the sun.

In this case we replace (12) with (13) and believe its content rather than imagining it.

Now let us consider a case in which we compare a fictional system and a real system. Consider again Giere's original example:

9. The positions and velocities of the earth-moon system are very similar to those of a two-particle Newtonian model with an inverse square force.

We imagine that the content of (9) obtains and assert something that is fictionally true within an extended imagined context involving the Newtonian model system and the

real earth-moon system. Analysis₁ can account for the intuition that (9) is genuinely true by embedding it within an extended fiction:

14. According to the extended fiction, the positions and velocities of the earth-moon system are very similar to those of a two-particle Newtonian model with an inverse square force.

Considering how a two-particle Newtonian model system is described in the model and how the positions and velocities of the earth-moon system really are, the positions and velocities of the earth-moon system are very similar to those of a two-particle Newtonian model. In this case, belief is the relevant attitude.

Here is one worry one may have with Analysis₁. Odenbaugh (2015) recently claimed that the fiction view of modelling advanced by Toon (2012) and Frigg (2010a) could not explain how scientists learn anything about the real world target because ‘whatever we learn about it, we learn about it in the make-believe’ (*ibid.*, 285). And even more seriously, he claims, models are explanatory and predictively accurate, yet ‘if modeling is a form of make-believe, then this scientific success is make-believe as well. The predictive and explanatory success due to modeling only occurs in a game’ (*ibid.*).

The account of model-world comparisons I just offered recognizes a fundamental and essential role to our imaginative engagement with fictions, yet it also explains how we can assume an external, non-participatory – and hence non-imaginative – perspective towards model-world comparisons. We explore and develop models in the imagination, from an internal or participatory perspective. And we originally compare models and targets from within an extended imagined context.

In this case, we are fully immersed in the game and we assume an attitude of imagination towards the relevant propositions. However, we can also exit the game and assume an external and descriptive perspective by quantifying over fictions and embedding the relevant propositions within a fictional operator. The attitude we assume towards these propositions is one of belief, rather than imagination, and the relevant statements can be evaluated for genuine truth in the real world. The belief that the new proposition is true in the world is not part of the game.

Now the skeptic could grant that when assessed from a real world perspective, model-world comparisons are not part of any game. They are genuine assertions. Yet, she could say, this is not enough. Scientific success requires that we get the claims *out of the fictional operator* to be able to say of the real earth-moon system – on the basis of the model – that, say:

15. The square of the orbital period of a planet is proportional to the cube of the semi-major axis of its orbit.

This, however, is a different claim. For one thing, it is not a comparison. It is a claim that can be made about the model system and within the model (in which case it will be fictionally true) or without the model (in which case it will be true only if prefixed by the fictional operator). But it can also be a *testable hypothesis* about the real system. In the latter case (15) is the outcome of the imaginative activity performed within the game that has been exported as a hypothesis about a real system. Doing this requires that we first assume that some relevant features of the model are shared by the real system in the imagination. Once we have produced the relevant model-world comparisons we can select and then export the hypotheses about the real

system. Such exports are one step removed from the imagination, but they could not have been achieved without going through the imagination first.

When scientists quantify over the relevant degrees of similarity we can apply Analysis₂ if we assume realism about mathematical entities and quantify over them. Consider (9) again. In this case, we can appeal to external quantification over the values describing the positions and velocities of the real earth and moon and of the imagined two-particle Newtonian system embedding internal quantification over the Newtonian model:

16. There are some values for positions x_1 and x_2 and velocities v_1 and v_2 , such that $x_1 \equiv x_2$, $v_1 \equiv v_2$, the earth-moon system has position x_1 and velocity v_1 , and according to the Newtonian model the two-particle model system has position x_2 and velocity v_2 .

Notice that this is essentially different from Frigg's (2010a) appeal to comparisons of uninstantiated properties. On Frigg's analysis model-world comparisons cannot be true because the properties attributed to the model system and compared with those of the real system are not instantiated. Analysis₂, however, delivers the right truth-conditions on the assumption that degrees of properties exist (they are mathematical entities) and according to the model they are instantiated by a certain model system. On this analysis, an utterance of (16) is a comparison between some mathematical entities (the values for positions x_1 and x_2 and velocities v_1 and v_2), when x_1 and v_1 are really instantiated by the earth-moon system and according to the Newtonian model x_2 and v_2 are instantiated by the two-particle model system. The ontological commitment to the real existence of mathematical entities and the according-to-the-model operator

deliver the right truth-conditions. Thus, the model-world relation of similarity is a relation that is grounded in the construction, development and exploration of model systems in the imagination. Learning with models is learning through the imagination.¹⁰

Acknowledgments

I would like to thank Greg Currie, Roman Frigg, Arnon Levy, and Alisa Bokulich for extremely helpful comments on earlier drafts. Thanks also to Adam Toon for a very helpful e-mail exchange. Financial support for this research was provided by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement No 654034.

References

- Berto, Francesco 2011. Modal Meinongianism and Fiction: The Best of Three Worlds. *Philosophical Studies* 152: 313-334.
- Bokulich, Alisa 2016. Fiction as a Vehicle for Truth. *The Monist*.
- 2012. Distinguishing Explanatory from Nonexplanatory Fictions. *Philosophy of Science* 79 (5): 725-737.
- 2009. Explanatory Fictions. In M. Suárez, ed., *Fictions in Science: Philosophical Essays on Modeling and Idealization*. New York: Routledge, 91-109.

¹⁰ For an in-depth study of the sort of imagination involved in model-based science (and thought-experimenting) see Salis and Frigg (forthcoming).

- 2008. Can Classical Structures Explain Quantum Phenomena? *British Journal for the Philosophy of Science* 59 (2): 217-35.
- Braun, D. 2005. Empty Names, Fictional Names, Mythical Names. *Noûs*, 39: 596-631.
- Currie, Gregory 1990. *The Nature of Fiction*. Cambridge University Press.
- 1988. Fictional Names. *Australian Journal of Philosophy* 66: 471-488.
- Frigg, Roman 2010a. Models and Fiction. *Synthese* 172/2: 251-268.
- 2010b. Fiction in Science. In John Woods, ed., *Fictions and Models: New Essays*, Munich: Philosophia Verlag, pp. 247-287.
- Giere, Ronald N. 1988. *Explaining science: a cognitive approach*. Chicago: The University of Chicago Press.
- 2004. How Models Are Used to Represent Reality." *Philosophy of Science* 71:742–52.
- 2009. Why Scientific Models Should Not Be Regarded as Works of Fiction. In Mauricio Suárez (ed.) *Fictions in Science. Philosophical Essays on Modelling and Idealization*, 248-58. London: Routledge.
- 2010. An Agent-Based Conception of Models and Scientific Representation. *Synthese* 172: 269-81.
- Godfrey-Smith, Peter 2006. The strategy of model-based science. *Biology and Philosophy* 21: 725-740.
- 2009. Models and fictions in science. *Philosophical Studies* 143: 101-116.
- Levy, Arnon 2015. Modeling without models. *Philosophical Studies* 172(3): 781-798.
- Odenbaugh, Jay 2015. Semblance or similarity? Reflections on Simulation and Similarity. *Biology & Philosophy* 30 (2): 277-291.

- Richard, Mark 2000. Semantic Pretense. In A. Everett and T. Hofweber, eds., Empty Names, Fiction and the Puzzles of Non-Existence. Stanford, CA: CSLI Publications.
- Sainsbury, Mark R. 2010. Fiction and Fictionalism. Routledge.
- Salis, Fiora 2013. Fictional Entities. Online Companion to Problems in Analytic Philosophy.
http://repositorio.ul.pt/bitstream/10451/10860/1/Fictional%20Entities3_Salis%20Fiora_Companion2013.pdf
- Salis, Fiora and Roman Frigg forthcoming. "Capturing the Scientific Imagination," in Peter Godfrey-Smith and Arnon Levy, eds., The Scientific Imagination, New York: Oxford University Press.
- Stanley, Jason 2001. Hermeneutic Fictionalism. In P.A. French and H.K. Wettstein (eds.), Figurative Language. Boston: Blackwell, 36-71.
- Thomasson, Amie 1999. Fiction and Metaphysics. New York: Cambridge
- Thomson-Jones, Martin 2010. Missing Systems and the Face Value Practice. Synthese 172(2): 283-299.
- Toon, Adam 2012. Models as Make-Believe: Imagination, Fiction, and Scientific Representation. Palgrave Macmillan.
- Van Inwagen, Peter 1977. Creatures of Fiction. American Philosophical Quarterly 14 (4): 299-308.
- Walton, Kendal 1990. Mimesis as make-believe. Cambridge, MA: Harvard University Press.
- Walton, Kendal 1993. Metaphor and prop oriented make-believe. European Journal of Philosophy 1: 39-57.
- Weisberg, Michael 2007. Who is a modeler? British Journal for the Philosophy of

Science 58 (2): 207-233.

--- 2013. *Simulation and similarity using models to understand the world*. New York: Oxford University Press.

Yablo, Stephen 1999. A Paradox of Existence. In A. Everitt and T. Hofweber, eds., *Empty Names, Fiction, and the Puzzle of Non-Existence*, Stanford: CSLI, pp. 275-312.

--- 2014. *Aboutness*. Princeton, NJ: Princeton University Press.

Zalta, Edward N. 1983. *Abstract Objects. An Introduction to Axiomatic Metaphysics*. Dordrecht: Reidel.

--- 1988. *Intensional Logic and the Metaphysics of Intentionality*. Cambridge: MIT Press.