# If models are make-believe, how can we learn from them?

Joe Roussos

16 May, 2017

It is a basic feature of science that we investigate real systems through models. In some cases, these models are real physical systems; more often, they are non-material, theoretical systems. Theoretical models involve objects that are not real—such as ideal gases and perfectly rational agents—and investigations of them involves understanding the properties and relations of these objects. This process is regarded by scientists as an extremely fruitful way of learning about the world. But since the objects in the model do not exist, it seems that claims about properties of model systems are, strictly speaking, false. Given this, we cannot make fruitful comparisons between models and the real world. And yet science does this regularly.

This is the core question of this paper: how can we learn from models? An answer to this question will have at least two parts: (1) an account of model-world comparisons, i.e. an account of when and how statements that compare models to the real world are true; and (2) an account of statements about the real world only, motivated by models but not referring to them. As should be clear from the above this is in part an exercise in the semantics of model talk, and in part an exercise giving an account of the metaphysics of model systems.

I will examine two linked answers: Roman Frigg's (2010b) account, and Fiora Salis' (2016) development on it. I will begin by describing Frigg's account of models as fictions, and the account that he gives for the truth-conditions of various kinds of statements about models and the real world. I will then look at the specific challenge of learning from models, presenting a challenge to Frigg due to Peter Godfrey-Smith. I argue that this response is less important than commentators, including Salis, have assumed. I will then examine Salis' proposal for the semantics of model-world comparisons and argue that it is easily incorporated into Frigg's account if we are concerned with Godfrey-Smith's objection.

## 1 Frigg's theory of models as fiction

The approach to scientific models that I will explore in this paper rests on an analogy with fiction. (Fiction here refers to works of fiction, in literature and other arts.) The purpose of the analogy is to make use of philosophical analyses of fiction to illuminate the problems of models, and in particular the problem of how we can learn true propositions about the world from the study of models.

Before introducing fiction, let us reflect on the role of models in science. Many real world systems are too complex to be investigated directly by scientists. In some areas we are constrained by scientific theory—we simply don't know how to analyse all of the interactions involved—while in other cases the issue is computational complexity—perhaps the system contains too many interacting parts to be tractable. In order to make progress, scientists instead address "toy," simplified versions of the system under investigation. These simplified versions typically leave out many aspects of the real system which are regarded as outside of the scope of the current investigation. Other aspects are neglected despite being relevant, because their effect is regarded as too small to matter, or because, despite being of non-negligible size, the calculations required are too difficult to accomplish with present techniques and resources. An example of the former might be the neglect of friction in the bearings of a pendulum; of the latter, the neglect of air resistance when studying a falling sphere. Scientists might also alter the nature or size of effects to make the calculation simpler or to focus on certain features of the system. These simplified versions are called (scientific) models.

In rare cases, scientific models are physical systems; material objects such as John Kendrew's plasticine model of myoglobin (discussed in Frigg and Nguyen, 2016), or the U.S. Army Corps of Engineers Bay Model, a scale model of the San Francisco Bay developed in the early 1950's to investigate the effects of building two proposed dams in the bay.[1] The majority are theoretical models that do not exist as physical systems, such as the Rutherford-Bohr model of the hydrogen atom. Such a model might be *described by* a series of written statements (Bohr's 'rules' for electron orbits), often accompanied by an equation (e.g., $L = n\hbar$, specifying the allowed angular momenta of orbiting electrons) and perhaps illustrated with a diagram. But the *model system* we are investigating when we use the Bohr model is not identical with any or all of these physically instantiated parts; it is what those descriptive elements *specify*.

What is the model system? Godfrey-Smith (2009) outlines three approaches to this question in the literature.

1. **Structuralism**: Giere (2009, 2004)[2] argues that model systems are abstract entities, that stand in relations "of similarity in certain respects and to certain degrees with concrete systems" (Salis, 2016, 3).[3]

[1]See http://www.spn.usace.army.mil/Missions/Recreation/Bay-Model-Visitor-Center/The-Bay-Model-Journey/History/

[2]There are two views on models in the literature called Structuralism, one due to Giere and discussed above, and the other due to Suppes (1969). I will not examine either in detail, nor comment on whether it is appropriate to use the same name for the two.

[3]Page numbers for Salis (2016) refer to the LSE online preprint.

2. **Indirect Fictionalism:** Frigg (2010b), and Frigg and Nguyen (2016), argue that models are hypothetical concrete objects, like characters or places in fictions. If they existed, they would be material objects with ordinary physical properties.

3. **Direct Fictionalism** Toon (2010) and Levy (2015) argue that there are no model systems; talk of them is a "psychological prop for organizing talk that achieved something quite different," that being the development of "approximately true descriptions of real systems" (Godfrey-Smith, 2009, 114).

This paper is an examination of the second approach. As I will not be comparing it to the third, I will simply call it Fictionalism.

### 1.1 Fiction as make-believe

Why think that models are like fictions? Salis and Frigg (2017) open their paper with a series of quotations from Maxwell, Einstein, and others inviting their readers to *imagine* a scenario or object, in the context of introducing a scientific model or thought experiment.[4] The subject of these imaginings are often non-existent objects like ideal gases or frictionless planes. Godfrey-Smith takes it that a "natural first description of these things is as fictions," as this is the common term we use to describe such non-existent "creatures of the imagination" which we want our audience to entertain for a time and a purpose (2009, 101). In the case of literature, the purpose is entertainment (and, perhaps, the investigation of some general truths about people); in science, the purpose is to learn about the real world.

What do we mean by 'imagination'? It is an attitude distinct from, say, belief; when we read a work of fiction, we don't believe the propositions it expresses, we *imagine* them. Imagination "involves more than just entertaining or considering or having in mind the propositions imagined" (Walton, 1990, 20). When one imagines a something, one is "*doing* something *with* [the] proposition one has in mind" (20). This, as Berberousse and Ludwig (2009) say, makes imagination "a powerful means of achieving understanding of various situations. As Currie and Ravenscroft emphasize, it enables us 'to project ourselves into another situation and to see, or think about, the world from another perspective' (Currie and Ravenscroft, 2002, 1). 'Think about the world' here is not just a figure of speech. Because imagination preserves inferential patterns of belief (Currie and Ravenscroft, 2002, 13), we may reason about what we imagine and draw conclusions, exactly in the same way as we do with out beliefs" (Berberousse and Ludwig, 2009, 58).

Frigg's account of models uses Walton's 1990 theory of fiction as an exercise in a particular kind of imagination, called make-believe. On this theory, fictional texts do not provide *descriptions* of nonexistent objects, but rather *prescriptions*, instructions for the imagination (Levy, 2015, 788). Frigg's (2010b) account of the theory is as follows. Walton defines 'make-believe' as human exercise in imagination aided by external objects

[4]Salis and Frigg argue that thought experiments and scientific models function in essentially the same way. I agree with this, but will not pursue this matter in this paper. The argument here does not depend on this stance.

called 'props.' An object becomes a prop due to the imposition of a rule or 'principle of generation,' prescribing *what is to be imagined* when the object is present. If someone imagines something because of a prop, he is "engaged in a game of make-believe" (2010b, 258). Some rules of generation are ad hoc, like when children playing make-believe declare that the floor is lava, and Walton calls these 'unauthorized' games. Other rules of generation are public and relatively stable, and the associated games called 'authorized'—an example here could be a published fictional story.

Props in authorized games are called 'representations'; a confusing usage as this is not representation as philosophers are used to it. P-representation, as Frigg calls Walton's usage, is not about relation to a target that is represented; instead, p-representations are things that possess the social function of serving as props in authorised games of make-believe. (The more common philosophical notion of representation of a target will be called *t-representation*.) *Harry Potter* (the series of books) p-represents a person called Harry because everyone who reads the book is invited to imagine that character, but t-represents no one. A hairbrush does not p-represent a microphone, no matter how enthusiastically a teenager sings into it, because this is not *authorised* make-believe.[5]

Within fictions, we naturally describe certain statements as true, using this form of words to distinguish between correct and incorrect descriptions of what happens in the fiction. Since these fictional happenings don't exist, this is at least in prima facie conflict with our standard usage of 'true.' Frigg therefore introduces a tripartite system of propositions about fictions, to which I add a fourth, with corresponding propositional attitudes and truth conditions.

1. *Intrafictional* propositions "are made within the fiction," and the attitude we take to them is imagination (rather than belief).

   (1) Harry Potter is a wizard.

   Read literally, statements like (1) are false, because there is no Harry Potter and so we need an account of *truth in the fiction*. In Walton's theory, "props generate fictional truths by virtue of their features and principles of generation" (259). Some such fictional truths are 'directly generated' by the prop, as they are part of what we are invited to imagine. Other fictional truths can be generated 'indirectly' by the application of rules of inference to direct truths. We are invited to imagine a magical England in *Harry Potter*; we can infer that there is a United States of America and that it likely has wizards too, although nothing of the sort is mentioned in the books. (In the terminology of this theory, one speaks of something 'being fictional' as an alternative to the locution 'is true in the fiction.')

2. *Metafictional* propositions are *about* the fiction rather than *within* it, and so "can be true or false in the same way as claims about chairs and tables" (262). We can therefore properly take the attitude of belief toward them.

[5]In (Frigg and Nguyen, 2016) the authors clarify that they regard Walton's theory as providing a successful account of a *specific genre* of fiction ("straightforward narration of the kind we find in late 19th-century novels"), and that their claim is that "model descriptions function like texts of this kind and that therefore [pretence theory] offers a good account of them" (240, fn13).

(2) According to *Harry Potter*, Harry Potter is a wizard.

Salis puts the truth conditions of (2) thus: "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 [the associated work of fiction,] and false otherwise" (2016, 7). Put another way, when we assert (2), we assert that "it is fictional in *Harry Potter* than Harry Potter is a wizard" which is straightforwardly true.

3. *Transfictional* propositions compare elements of one fiction with another, or compare elements of a fiction to elements of the real world. They require some imagination, so it is unclear to begin with what attitude we should take toward them. How can we account for the truth-conditions of statements like (3)?

   (3) My friend Luke was a worse school student than Harry Potter.

   Frigg's (2010b) answer is that "truth conditions for transfictional statements [...] come down to truth conditions for comparative statements between properties" rather than between the real object Luke and the nonexistent object Harry Potter (263). Salis disagrees. This will be developed in detail below.

4. *Extrafictional* propositions are not about the fiction at all; they concern the real world but are *justified by* the fiction. In literature an example might be a claim about human nature ("love conquers all") that one has learned from reading a novel. These will have straightforward truth-conditions (they concern the real world) and the proper attitude toward them will be belief. But we will want to say something about how we come to be justified in believing them, based on the fiction.

### 1.2 Models as fictions

So much for fictions, what about models? In analogy with the 'fictional world' of, say, *Harry Potter*, we have the 'model world,' which contains the model system (as specified by the model description, which itself comprises many of the primary fictional truths) and various rules of inference that allow us to fill out the model world as we need to (thereby generating indirect fictional truths). But scientific models are more complex than literary fictions, in that they involve two steps: first, "introducing a p-representation specifying a hypothetical object and, second, claiming that this imagined object t-represents the relevant target system" (Frigg, 2010b, 264).

Note that no claim is being made that the *purposes* of scientific models are the same as those of fictions. Nor is there any intention to degrade the standing of science by comparison to fiction. Nor is the distinction between fiction and non-fiction being denied. Rather what is asserted is that the commonality between models and literary fictions is in the attitude one is to take to their contents: imagination. This, rather than just truth or falsity, distinguishes fiction from non-fiction. The status of science is not challenged because, as this essay aims to show, the truths that we extract from models are true (if they are) in the same way that ordinary empirical claims are, despite models being fictions. This therefore means we need not worry that the Fictionalist account has anything radical to say about the purpose of science.

We can now begin the analysis of various statements about model-systems and their properties. Intrafictional truths are statements within the model, which are 'true in the model' if they are part of the prescribed imagining of the model. "For instance, 'the solar system is stable' is true in the Newtonian model of the solar system iff the description of the system together with the laws and principles assumed to hold in the system [...] imply that this is the case" (262). When scientists work within the framework of a model, they will make many such intrafictional claims. The correct attitude to the propositions these express, says Frigg, is imagination. We aren't intended to believe them, because they aren't about the real world.

This strikes me as a very plausible account of the practice of using models, at least in physics. Physicists are often explicit that the systems they are investigating do not exist. Consider this line from Griffiths' classic quantum mechanics textbook: "The paradigm for a classical harmonic oscillator is a mass $m$ attached to a spring of force constant $k$. ...Of course, *there's no such thing as a perfect simple harmonic oscillator*—if you stretch it too far the spring is going to break, and typically Hooke's law fails long before that point is reached" (my emphasis, Griffiths, 1995, 31). Scientists consider, entertain, and imagine model systems, but they don't *believe* the description involved apply to any real system. This is occasionally highlighted by the use of metafictional statements *about* the model in question, such as 'In the Newtonian model of the solar system, the solar system is stable.' (These are evaluated exactly as in the fiction case.)

Providing an account of transfictional statements about models, and extrafictional statements derived from them, is more complicated, and will begin to answer the question of learning from models. Here is Frigg's explicit cashing out of comparative statements between a model and reality:

> When I say that the population of rabbit in a certain ecosystem behaves very much like the population of the predator-prey model, what I assert is that these population possess certain relevant properties which are similar in relevant respects. [...] the statement making this comparison is true iff the statement comparing the properties with each other is true. (Frigg, 2010b, 263-4)

The picture appears to be this: the model 'has' a property (say, its 'population growth rate') which we can compare with a property of the real world population of rabbits (viz. the real rate of population growth). Our comparisons don't rely on comparisons between nonexistent model-rabbits and real rabbit, but rather between the properties imagined as part of the model-system and those truly instantiated by the real rabbits.

But how does this comparison of properties occur? To understand this we need more information on how t-representation works. If we can spell out how models t-represent real world systems, we will better understand how to extract meaning from comparisons between the two, and will in the process derive an account of extrafictional learning. To fill out the notion of t-representation, Frigg (2010a) suggests an analogy with maps, specifically the fact that maps are accompanied by *keys* which specify "how facts about the map translate into facts about the city" (126). T-representation is then schematised by the following two conditions:

$X$ t-represents $Y$ iff:

(R1) $X$ denotes $Y$.

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

The first condition establishes reference, that the representation is *of* the target system. The second condition explains which parts and properties of the model-system are relevant (not all are) and how they are intended to map onto the properties of the real system. This is intended, says Frigg, as a "*general form* of an account of t-representation, which needs to be concretized in every particular instance of a t-representation" (128). By this he means that maps, graphs, paintings, and scientific models will all achieve their t-representation in different ways; but in each instance this will include (R1) a way in which the representation denotes the represented, and (R2) an (implicit or explicit) set of rules for which parts of the representation are salient and how they correspond to the represented.

One important example of a key discussed by Frigg is the *ideal limit*.

> This happens when we model particles as point masses, strings as massless, planets as spherical, and surfaces as frictionless. Two things are needed to render such idealizations benign: experimental refinements and convergence [...] First, there must be the possibility of in principle refining actual systems in a way that they are made to approach the postulated limit (that is, we don't actually have to produce these systems; what matters is that we in principle could produce them). [...] Second, this sequence has to behave 'correctly': the closer the properties of a system come to the ideal limit, the closer its behaviour has to come to the behaviour in the limit. (Frigg, 2010a, 131)

We begin an exercise in scientific modelling with an aim, e.g., determining the orbit of the earth around the sun. This establishes the target for the model. In the formulation of the model, we make a set of assumptions about the properties relevant to our aim. We make use of scientific theories to provide us with the equations held to govern the system, under those assumptions, e.g., Newton's law of gravitation. Further simplifying assumptions are now often required; idealisations that simplify the calculation in the manner described above. Having completed this process, we have the model description and when we imagine what it described we have the model system. In our example, we might imagine two perfect spheres, each of homogeneous density, one much lighter and orbiting the heavier body.

In some cases, like ours, there will be an explicit model equation specifying the model's structure. It acts as a principle of generation, a rule allowing us to make inferences to properties of interest such as the orbit of a smaller body around a larger body. We now connect back to the target system, beginning with denotation (R1): the lighter body corresponds to the earth, the heavier to the sun. For (R2), the relevant key involves a series of ideal limits (e.g., the density and shape of the balls, the absence of other masses in the universe, etc.). Once we understand these limits, we can translate facts about the model system into claims about the target system. "For instance, calculations reveal that the model earth moves on an ellipse, and given that the model system is an ideal limit of the target we can infer that real earth moves on a trajectory that is almost an ellipse" (Frigg, 2010a, 135).

That final statement is clearly extrafictional: we investigate the model system, and once we have learned something about its properties, we *impute* certain properties of the real system—in this case, a nearly elliptical orbit. These are now empirical hypotheses, they can be tested and found true or false. The modelling process provides some justification for believing them (prior to testing) under certain conditions; for example if we have confidence that the idealisations made in the model lead to reliable results of this sort. In practice, this will depend on previous hypotheses generated from this same model having turned out to be true (or true enough for the current purposes).

We can also make explicit model–world comparisons. With (R1) and (R2) in place, we know which object in the model corresponds to which object in the real world. The model equation(s) allow us to calculate the values of properties (positions, momenta) that the key tells us are intended to represent the properties of the real earth. We can then compare these calculated values to measured values of the real earth. But we can also make non-numerical comparisons. We might say for instance that "the perfect spherical nature of the model earth differs from the shape of the real earth." In so doing we are using the imaginative process to latch onto a property, being spherical, and comparing it to an actually instantiated property, that of the shape of the earth.

## 2 Criticism of Frigg's model

Frigg's model has been subject to one major criticism, articulated by Peter Godfrey-Smith (2009) and later endorsed by Levy (2015) and Salis (2016). It is worth quoting Godrey-Smith in full:

> Frigg treats non-existent objects as problematic—something to be avoided. He thinks properties, however, can be used in an explanation with less anxiety. It is not that properties are entirely unmysterious, but the problems they bring must also be confronted in philosophical contexts that have nothing to do with fiction. I am not so sure that there is such a difference between the candidates. When dealing with fictional models, many of the properties that are being introduced will be uninstantiated ones. These do seem to raise special problems of the same kind as those seen with fictional objects. (Godfrey-Smith, 2009, 113)

As I read it, this is a criticism of dialectical strategy. Frigg thinks that it is a recommending feature of his theory that it does not require non-existent objects, because this allows the theory to avoid an entire arena of debate particular to (or centred on) fiction.[6] Therefore, he proposes that we learn about models through their properties. But,

[6] Frigg does not discuss this in depth, merely saying that "the metaphysics of fictional entities is an issue fraught with controversy. . . For this reason we need to know what kind of commitments we incur when we want to understand model systems along the lines of fiction, and how these commitments, if any, can be justified" (Frigg, 2010b, 257).

say critics, models regularly involve *uninstantiated properties*, a subject of thorny metaphysical debate. At the heart of this debate is the *principle of instantiation* which says that all properties are instantiated; there are no uninstantiated properties (see Orilia and Swoyer, 2016). The problem being that there are no infinitely malleable springs, or perfectly spherical planets as referred to in models; indeed, there are no infinitely malleable or perfectly spherical objects *at all*, and so, being uninstantiated, those properties do not exist.[7]

Salis and Levy write as though they are endorsing Godfrey-Smith's criticism but I think they in fact express a (common) different criticism. Here is Salis:

> 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. (Salis, 2016, 20).

This passage follows an earlier passage in which Salis says that on an object-dependent view, the claim "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" is false, because the (abstract object that on this view is) the Newtonian two-particle system cannot have any positions or velocities (2016, 15). It therefore seems to me that Salis' point is that, because the model-system does not exist it (is merely imaginary) it does not *itself* instantiate properties, and therefore "model–world comparisons are still false."

Levy expresses what he says is the same concern slightly differently: for him, the issue is that comparisons require some notion of difference or resemblance. "Resemblance, whatever exactly it comes down to, is a relation between objects (or events, or at any rate, *bona fide* things). But on the make-believe view, the model is not a thing in any robust sense—it is only a set of prescriptions for the imaginations of scientists. How then are models to be compared with targets (or, for that matter, with one another)?" (Levy, 2015, 789) Again, the issue is the non-existence of the *model system*, which for Levy makes model–model and model–world comparisons impossible.

So Frigg faces two distinct criticisms:

(C1) Frigg's dialectical strategy fails because it embroils him in an equally difficult debate about properties (e.g., the Scholastic debate about the existence conditions of universals).

(C2) Model systems are not real, there are merely imaginary, and therefore (Salis) all model comparisons are false, or (Levy) impossible.

[7] I understand the principle of instantiation as follows. All genuine properties are instantiated, instantiation is a necessary condition for being called a "property." When considering descriptions which have the superficial appearance of being properties but are *not instantiated*, like infinite malleability, we therefore say that they 'are not properties,' or that they 'do not exist.'

Let us look at each in turn. First, how thorny a problem *is* (C1) for Frigg? Recall the *principle of instantiation*: there are no uninstantiated properties. So if we are ascribing an uninstantiated property to an object, this ascription cannot be true for there simply is no such *property*. In the case of a material object (for example a physical model) the problem would not exist: if the object had the property, then it would instantiate that property. But as theoretical model systems are imaginary, they cannot instantiate properties. Note first that the principle of instantiation is not universally accepted. If we are willing to adopt Platonism about properties (by which I mean any theory that has properties existing independently of instantiation), then there is no problem.

Suppose we are not Platonists. Orilia and Swoyer (2016) distinguishes between two forms of the principle of instantiation: weak and strong. On the strong principle, properties must be physically instantiated (by things that exist in space and time); while on the weak version they can be instantiated by abstract objects. If we work with the weak version, then so long as one is willing to admit mathematical objects into one's ontology, the objection dissolves. There are no physical one-dimensional strings, or infinite populations, but there are one-dimensional mathematical objects, infinite sets, and so forth. So these properties are instantiated, and there is no problem. (Or rather, there are difficulties in the analysis of properties, but these are not of a kind with the problems of non-existent objects, which Frigg's dialectical strategy was targeted at.)

On the strong principle of instantiation, these properties are uninstantiated. But this view seems problematic for a philosopher of science to hold. If properties must be physically instantiated, then their existence is contingent. There are two options for the existence conditions of contingent properties: either they exist only while they are instantiated, or if they are instantiated at any time, then they exist at all times. On the former account, we encounter counterintuitive results with laws of nature. Laws are or express relations among properties. But if properties come in and out of existence then so too do laws. This contradicts their lawlike nature; the intuition we have that we are seeking robust generalisations that represent more than merely contingent states of affairs. If properties exist if they were ever instantiated, then we face an epistemic problem: how do we know whether the properties we call upon in our laws exist? This isn't subject to empirical investigation, as the existence condition spans all of time. These brief remarks are not intended to be definitive in this complex metaphysical debate; they merely gesture at reasons that, at the level of dialectical strategy, philosophers of science may not wish to endorse Godfrey-Smith's critique. If none of his interlocutors would consider the views in the relevant contrast class plausible, this does not seem like a major issue for Frigg's account.

We can now turn to (C2) and ask whether model–model and model–world comparisons are (a) possible, and (b) if so, whether Frigg's theory can accommodate non-trivial truth-conditions for them. First, let us note that this is an issue for transfictional comparisons, but not for extrafictional claims. Salis acknowledges this, characterising an extrafictional claim as "a *testable hypothesis* about the real system... [It] is the outcome of the imaginative activity performed within the game that has been exported as a hypothesis about the real system" (2016, 27). These hypotheses, as Frigg and Nguyen point out, are "of the form 'target $T$ has property $Q$' ... [which] are standard attributive claims rather than comparisons, and as such they raise no problems having to do with fiction" (Frigg and Nguyen, 2016, 239).

In the absence of further motivation from Salis or Levy I find it hard to comprehend why the non-existence of the model system blocks property comparisons. Recall my ideal limit example. In that scenario we imagine a perfectly spherical planet. We then say "the real earth is not perfectly spherical but rather oblate." How could this possibly be problematic? I fix on the relevant property through the process of make-believe: I engage with the model description, follow the rules of generation outlined by the equations, and come to understand that I am being asked to imagine a perfect sphere. Given that I am not a believer in the strong instantiation principle, I hold that the property of being perfectly spherical exists. I therefore feel capable of performing the required comparison.

Perhaps the objectors would reply that I cannot compare these two planets in the same way that I can compare two physical objects: one a sphere, the other oblate. There I could place them next to one another, compare their shapes visually, and so forth. This is true. But it is not part of the account that we investigate the properties of model systems *in the same way* that we do with physical systems. Here is a simpler situation. Imagine a red cup. Now locate a blue object somewhere in your field of vision. Do you feel incapable of saying "this blue object is not the same colour as the cup I imagined"? Does your inability to precisely compare the hue of the real object and the cup block this comparison? If not, it seems you have no grounds for worrying about the comparison of model properties with real world properties.

## 3 Salis on transfictional comparisons

Fiora Salis (2016) takes it that these comparisons are not possible, or at least that they raise issues serious enough to warrant a work-around. She therefore develops a sematics of transfictional comparisons that does not require the comparison of imagined and real properties. I will now examine this account and argue that it is easy to incorporate into Frigg's account, should we be so minded.

Salis provides two potential analyses of transfictional comparisons, which she calls $\text{Analysis}_1$ and $\text{Analysis}_2$. I will focus on $\text{Analysis}_2$, and discard these labels as superfluous. The aim is to transform any transfictional comparison into a (true) metafictional claim. Salis' analysis has two parts. First, comparisons between properties are "genuine comparisons between *degrees* of properties which are mathematical entities that can be individuated on a scale of measurement" (Salis, 2016, 12). Second, comparison statements are replaced with statements in which the claims about fictions are embedded within an in-the-fiction operator. The two are combined into a statement in which we externally quantify over degrees of properties, and embed internal quantification over the relevant fictions.

Examples will serve to clarify. Let us begin with a model-model comparison: suppose we want to compare *Games of Thrones* character Jaime Lannister to the Arthurian knight, Sir Lancelot, by saying

(4) Jaime Lannister's love for his Queen is a greater crime than Lancelot's love for his Queen.

Then under Salis' analysis the comparative clause claims that there are degrees of criminality $i$ and $j$, such that $i > j$, Jaime's love is a crime to degree $i$ and Lancelot's to degree $j$. At this stage we still have a truth-conditional problem: Jaime and Lancelot do not exist, and so the claims about them having properties to degree $i$ and $j$ respectively will come out false. Therefore, we need to embed each part of (4) within an in-the-fiction operator, to make the statements about each character fictional ascriptions rather than attempted predications.

(5) There are some degrees of criminatilty, $i, j$, such that $i > j$, and according to *Game of Thrones* Jaime's love for his Queen is criminal to degree $i$, and according to Arthurian legend Lancelot's love for his Queen is criminal to degree $j$.

This creates a combined fiction containing both Jaime Lannister and Sir Lancelot. In this unauthorised game of make-believe, the fictional truth that Jaime loves his Queen, who is also his sister, and the fictional fact that Lancelot loves his queen, together with a background assumption that incestuous adultery is a greater crime than mere adultery (present in the real world, though also, I assume, in the fictions), entail that Jaime's crime is greater than Lancelot's. This is fictional, and (4) is a metafictional statement, which is therefore true and which we can believe.

In the case of models, Salis considers the following comparison of a result from the Rutherford-Bohr model with a result from the Newtownian model of the solar system.

(6) Electrons orbit about the nucleus just like planets orbit around the sun.

As above, we want to introduce some degrees of properties, and fictionally ascribe these to each model-system, in statements that are within the scope of an in-the-fiction operator. This will transform this transfictional comparison into a (true) metafictional claim:

(7) There are some values for masses $m_1 \gg m_2$, $m_3 \gg m_4$, and orbital eccentricities $e_1 \approx e_2$ and, according to the Rutherford-Bohr model of the atom electrons orbit about the nucleus governed by an inverse-square law, with the mass of the electron $m_1$ much smaller than the mass of the nucleus $m_2$, and the resulting orbit is circular ($e_1 \approx 0$); and according to the Newtownian model of the solar system planets orbit around the sun governed by an inverse-square law, with the mass of the earth $m_3$ much smaller than the mass of the sun $m_4$, and the resulting orbit is nearly circular ($e_2 \ll 1$).

Now consider a model–world comparison.

(8) 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 can carry out the same procedure as above (this time simpler, as there is only one fictional operator corresponding to the Newtonian model). The result is a "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" (Salis, 2016, 28). Correctly analysed, an utterance of (8) is genuinely true, rather than fictional. The cost is the introduction of the in-the-fiction operator, and a commitment to the existence of mathematical entities.

The difference with Frigg's proposed semantics of transfictional statements is this. There, the "properties attributed to the model system and compared with those of the real system are not instantiated." So, in Salis' reading, the comparison statements between them and real world properties are automatically false. Salis' analysis, 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" (2016, 28). It seems to me if one credited the problem, one could straightforwardly incorporate this semantic analysis into Frigg's system, by merely replacing his prescription for the treatment of transfictional statements with the above. I do not think the problem is a major one for Frigg's account, but I do think that Salis' semantics make for a neater overall system, aligning more directly with intra- and metafictional statements.

## 4 Conclusion

Fictionalism provides an account for what it is we're doing when we work with models in science. We are engaged in a game of make-believe, subject to specific rules derived from theory. Model systems are like characters in fiction; if they were real, they would be physical objects with properties. We access them through the imagination, and are able to extract knowledge from them in two ways. First, via transfictional statements comparing model systems to real systems. These are to be understood under Salis' analysis: they make existential claims about certain mathematical properties, fictionally attribute them to model systems, and then require a comparison between the fictionally attributed properties and those of real world systems. Second, we can generate empirical hypotheses about the real world from our investigation of the model system. Much more remains to be said about the conditions under which we are justified in crediting such hypotheses absent empirical investigation. For now we can simply conclude that we have in Fictionalism an illuminating account of the enterprise of working with and learning from models.

## References


Berberousse, A. and P. Ludwig (2009). Models as Fictions. In M. Suarez (Ed.), *Fictions in Science*, Chapter 4, pp. 56–73. London: Routledge.

Currie, G. and I. Ravenscroft (2002). *Recreative minds*. Oxford: Clarendon Press.

Frigg, R. (2010a). Fiction and Scientific Representation. In R. Frigg and M. C. Hunter (Eds.), *Beyond Mimesis and Convention*, Chapter 6, pp. 97–138. Springer Netherlands.

Frigg, R. (2010b). Models and Fiction. *Synthese 172*(2), 251–268.

Frigg, R. and J. Nguyen (2016). The fiction view of models reloaded. *Monist 99*(3), 225–242.

Giere, R. N. (2004). How models are used to represent reality. *Philosophy of Science 71*, 742–52.

Giere, R. N. (2009). Why scientific models should not be regarded as works of fiction. In M. Suarez (Ed.), *Fictions in Science*, Chapter 14, pp. 248–258. London: Routledge.

Godfrey-Smith, P. (2009). Models and fictions in science. *Philosophical Studies 143*(1), 101–116.

Griffiths, D. J. (1995). *Introduction to Quantum Mechanics*. Prentice Hall.

Levy, A. (2015). Modeling without models. *Philosophical Studies 172*(3), 781–798.

Orilia, F. and C. Swoyer (2016). Properties. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Win 2016 ed.). Metaphysics Research Lab, Stanford University.

Salis, F. (2016). The nature of model-world comparisons. *Monist 99*(3), 243–259.

Salis, F. and R. Frigg (2017). Capturing the Scientific Imagination. In P. Godfrey-Smith and A. Levy (Eds.), *The Scientific Imagination*. Oxford: Oxford University Press. (forthcoming).

Suppes, P. (1969). A comparison of the meaning and uses of models in mathematics and the empirical science. In P. Suppes (Ed.), *Studies in the methodology and foundations of science*, pp. 10–23. Dordrecht: Reidel.

Toon, A. (2010). The ontology of theoretical modelling: Models as make-believe. *Synthese 172*(2), 301–315.

Walton, K. L. (1990). *Mimesis as make-believe: On the foundations of the representational arts*. Cambridge, MA: Harvard University Press.