

Learning through the Scientific Imagination

Fiora Salis

University of York

Abstract

Theoretical models are widely held as sources of knowledge of reality. Imagination is vital to their development and to the generation of plausible hypotheses about reality. But how can imagination, which is typically held to be completely free, effectively instruct us about reality? In this paper I argue that the key to answering this question is in constrained uses of imagination. More specifically, I identify make-believe as the right notion of imagination at work in modelling. I propose the first overarching taxonomy of types of constraints on scientific imagination that enables knowledge of reality. And I identify two main kinds of knowledge enabled by models, knowledge of the imaginary scenario specified by models and knowledge of reality.

Keywords: Scientific models, Imagination, Make-believe, Counterfactual imagination, Knowledge.

1. Introduction

How do we learn about reality through scientific models? Answering this question requires distinguishing between two main kinds of models, material and theoretical. Material models are physical objects that serve as representations of physical systems. Theoretical models are mathematical models that do not exist as physical objects and for this reason are sometimes called ‘fictional’. Morgan (1999) originally argued that learning with models involves two steps, model building and model manipulation. Frigg and Hartmann (2018) notice that material models are used in experimental contexts and do not raise any special problems beyond general questions about learning through experimentation. Fictional models, however, do raise serious concerns. What are the constraints on model building and model manipulation in fictional models? Answering this question requires that we recognise the crucial role of imagination in fictional models. Learning through fictional models requires imagination and the value of scientific imagination depends on its ability to produce valuable (that is, potentially true) hypotheses.

Consider a simple but paradigmatic example, Maxwell's thermodynamic ideal gas model, which represents a gas as a large number of particles bouncing against each other and against the walls of a closed container. The model is usually identified with the following equation:

$$pV = nRT.$$

The equation per se is not a model of anything unless it is used under an interpretation. In this case, p is pressure, V is volume, n is the number of moles, R is the gas constant, and T is temperature. To facilitate mathematical treatment, the model makes certain simplifying assumptions. It assumes that the gas is composed of molecules construed as point particles having no volume in and of themselves, exerting no intermolecular forces, and bouncing against each other and against the walls of the container in elastic collisions that do not involve any conversion of kinetic energy into other forms of energy. Of course, there are no gases that are composed of such idealised molecules. Real gases are composed of molecules that have some finite volume, that exert intermolecular forces and collide in non-elastic ways. The ideal gas model describes an imaginary gas composed of imaginary particles interacting under imaginary conditions. Nevertheless, the model provides a useful approximation of the behaviour of many real gases under temperatures that are near room temperature and pressures that are near atmospheric pressure.

Philosophers usually recognise that imagination has an important role in modelling. Cartwright understands modelling as offering "descriptions of imaginary situations or systems" (2010: 22). Godfrey-Smith suggests we "take at face value the fact that modelers often take themselves to be describing imaginary biological populations, imaginary neural networks, or imaginary economies" (2006: 735) and sees modelling as involving an "act of imagination" (2009: 47). Harré sees models as things that are "imagined" (1988: 121). Sugden regards models as "imaginary" worlds (2009: 5). Weisberg discussing the Lotka-Volterra model of predator-prey interaction reports that Volterra "imagined a simple biological system" (2007: 208) and further recognises that "[m]odelers often speak about their work as if they were imagining systems" (2013: 48). Frigg (2010), Levy (2015), Salis (2019; 2020a), Salis and Frigg (2020), and Toon (2012) present analyses that place acts of imagination at the heart of modelling.

When it comes to explaining how models enable knowledge of reality, however, standard explanations dismiss uses of imagination in modelling as ill-suited to scientific reasoning. Notwithstanding their differences, these accounts agree in connecting learning with the representational function of models (Giere 1988; Knuuttila and Voutilainen 2003; Mäki 1992, 2005; Suárez 2004; Swoyer 1991; Weisberg 2007, 2013). On these views, a model description (the mathematical equation and linguistic assumptions of the ideal gas model) specify a simplified surrogate (the idealised gas) of a real system (some real gas). The surrogate is called model system and the real system is called target system. A model system is interpreted as a symbol representing a target in ways that enable the generation of plausible hypotheses based on a relation of similarity with the target, where similarity is usually understood as the sharing of certain properties in some respects and to some degree.

While these standard proposals advance many important ideas, they do not satisfy one key theoretical requirement that Frigg (2010) calls *naturalism*. According to naturalism, any account of how scientists learn with models should be able

to explain scientific practice, namely it should explain how scientists construct models and how they reason with them. This is what Thomson-Jones (2010) calls the face-value practice of modelling. Scientists present model-descriptions that specify model-systems as objects of study. Model descriptions involve the attribution of properties that only concrete objects can have, yet there are no objects instantiating these properties. Scientists think and talk as if there were such concrete systems having such and such properties, yet they are aware that there are none. They merely imagine that there are systems having such and such properties.

So, modelling crucially relies on imagination. Yet, standard accounts do not offer any explanation of how knowledge of reality is obtained through imagination. The result is a poor understanding of the epistemic role of imagination in modelling. Pre-theoretically, imagination is often thought of as completely free and unconstrained. In this vein, many think of imagination as a means to escape reality, as when we engage in daydreams and fantasies that provide diversion and create new things that depart from reality. Pessimists about our ability to gain knowledge through imagination emphasise the freedom of imagination (Descartes 1985; Norton 1991; Spaulding 2016). There is, however, another pre-theoretical notion of imagination as a means to learn about reality, as when we engage in problem solving, mindreading, thought experimenting, counterfactual reasoning and, of course, scientific modelling. The key to this second notion is the idea that imagination can be constrained in ways that effectively enable knowledge of reality.

In this paper I shed new light on this issue by developing a new notion of constrained imagination that is motivated by the face-value practice of scientists and by the recognition of the importance of scientific cognition involving imagination. In Section 2, I start by identifying two main varieties of imagination that are currently deemed crucial to scientific models, counterfactual imagination (Godfrey-Smith 2020) and make-believe (Salis and Frigg 2020). In Section 3, I argue that standard analyses of counterfactual imagination in modelling raise important issues that deserve further theoretical development. In section 4, I identify make-believe as a more suitable option and explain its role in model building and model development. In Section 5, I put forward a taxonomy of types of constraints operating on imagination in modelling based on contemporary literature in cognitive science and philosophy. Finally, in Section 6, I draw some conclusions.

2. Imagination

What sort of imagination is involved in models? Imagination is ordinarily construed as mental imagery, which is an ability to form a sensory-like representation of something (real or non-existent) in any sensory modality (imagining seeing, imagining hearing, imagining smelling, imagining touching, imagining tasting). The most common variety is visual imagery, which is often referred to as seeing in the mind's eye, imagining seeing or visualising. Scientists often appeal to this pre-theoretical notion in introspective reports and descriptions of activities that were key to the generation of new ideas. In the 19th century, Michael Faraday contributed to the foundations of classical electromagnetic theory by imagining invisible lines of force as narrow tubes curving through space (Tyndall 1868).

Starting from this picture, James Clerk Maxwell studied lines of force by producing a series of mechanical models of the ether, which led to his famous set of equations (Maxwell 1965). In the same century, August Kekulé discovered the structure of the molecule of benzene after a daydream in which he saw a snake biting its own tail (Findley 1948). These and similar cases led to the widespread recognition of the key role of imagery in scientific discovery, conceptual change and innovation (Magnani 2009; Nersessian 2008, 2009).

Whether imagery has a key epistemic role in modelling, however, is currently disputed. In particular, Salis and Frigg (2020) emphasise that mental images are neither necessary nor sufficient to scientific modelling. For example, the ideal gas model requires imagining that the gas be composed of point particles having no volume in and of themselves and bouncing against each other in elastic collisions. These imaginings involve certain theoretical concepts (point particle, volume, elastic collision) and relations within the imaginary scenario described by the model. Whether they are accompanied by mental images or not seems to be irrelevant to the epistemic function of models.

In fact, another notion of imagination has gained traction in the contemporary philosophical literature on scientific modelling, that of propositional imagination. This is an ability to entertain a proposition without any commitment to its truth, with or without forming a mental image. This somewhat minimal notion of imagination, which is akin to a notion of acceptance, has been specified in two main varieties that are deemed crucial to the modelling practice, counterfactual imagination and make-believe.¹

3. Counterfactual Imagination

Godfrey-Smith (2020) recognises the key role of conditional thinking and, in particular, the counterfactual imagination in modelling. Conditionals are statements of the form *if A then C*. A counterfactual conditional is a subjunctive conditional where the antecedent is known or assumed to be false, or $A \Box \rightarrow C$. For example, one might imagine that if Hillary Clinton had won the elections in 2016 (counterfactual antecedent), then the US would have led a coordinated effort to combat COVID-19 with allies in Europe, Asia and the Americas. Godfrey-Smith notices that counterfactual conditionals in modelling often involve generalisations such as *if there were a system like this, it would do that*, or $M \Box \rightarrow C$, where the antecedent M stands for the model assumptions and the consequent C stands for the consequence that follows from M . For example, if there were a gas having these and these features, then it would behave like this (ideal gas model); or, if there were two celestial bodies having features F , then they would do that (sun-earth model). The antecedents in these conditionals are assumed to be false. Scientists know that they are never realised in the actual world.

Implicit criteria for how imagination is constrained in counterfactual reasoning have been offered by the influential analyses of counterfactuals put forward by Stalnaker (1968) and Lewis (1973). The leading idea of these analyses is that a counterfactual claim is true in the closest possible world where the antecedent is

¹ Another important notion of propositional imagination is that of supposition (Arcangeli 2018; Nichols 2006), which plays an important role in Sorensen's (1992) account of scientific thought experiments. There are, however, no accounts of modelling in terms of supposition.

true and the consequent is also true. By ‘closest possible world’ we mean closest to the actual world—or reality. Hence, closeness—reality orientation or similarity—is the key constraint on imagination that emerges from these analyses. When engaging in counterfactual reasoning, we select an antecedent A that is contrary to some relevant fact in the actual world and then draw a consequence C in the A -worlds that are closest to the actual world. However, the antecedent A selects a set of possible worlds (the A -worlds), not all of which are relevant for the assessment of the counterfactual conditional. The A -world that is closest to reality is the one that determines its truth. When one ponders what would have happened if Hillary Clinton had won the elections in 2016, one considers how things would have been in a world that is just like the real world apart from the election of Hillary Clinton in 2016.

There is one general challenge for these analyses, and three specific issues concerning their application to modelling. The general challenge concerns the details of the notion of closeness, which remains insufficiently characterised. Stalnaker appeals to the “intuitive idea that the nearest, or least different, world in which antecedent is true is the one that should be selected” (1981: 88), but does not provide any explanation of what ‘least different’ means. And Lewis (1973) assumes a primitive notion of similarity of worlds, which “leaves the notion of similarity unconstrained and mysterious” (Arlo-Costa 2019: Sect. 6.1).

The three more specific challenges for an application of these analyses to models are posed by completeness, epistemic access, and intersubjective access.

Salis and Frigg (2020: 43) notice that possible worlds are complete, yet scientific models cannot be said to be complete in the same way. What completeness means is open to interpretation. However, they notice that there is an intuitive link between completeness and the principle of Excluded Middle (EM). According to EM, for any proposition p it is the case that *either* p or *not- p* holds. Models are not complete in this sense because there are many propositions that are neither true nor false in models. For example, the proposition that Mont Blanc is the tallest mountain in Europe is neither true nor false in the ideal gas model. However, if possible worlds are complete, the closest possible world in which M is true is one in which this claim is true even though it describes matters of fact that have nothing to do with the model. On this analysis, the counterfactual “if a gas were composed of point particles exerting no intermolecular forces, then Mont Blanc would be the tallest mountain in Europe” would come out true. The truth value of this counterfactual, however, should be indeterminate. The world of the model does not satisfy EM and is not complete in the same way in which possible worlds are supposed to be complete.

Stalnaker’s semantic analysis of counterfactuals does not allow for this kind of indeterminacy because it accepts the principle of Conditional Excluded Middle (CEM). According to this principle, either $M \Box \rightarrow C$ is true or $M \Box \rightarrow \neg C$ is true. Stalnaker’s semantics uses a selection function that picks a unique closest possible world where C is either true or false and, hence, either $M \Box \rightarrow C$ or $M \Box \rightarrow \neg C$ holds. In contrast, Lewis’s semantics deploys a relation of comparative similarity that defines a weak total ordering of all possible worlds with respect to each possible world (what he calls ‘a system of spheres’). On this proposal, when $M \Box \rightarrow C$ is true, C is true in all the closest M -worlds. However, when C is true only in some of the M -worlds but not in others, CEM fails because neither $M \Box \rightarrow C$ nor $M \Box \rightarrow \neg C$ holds. This allows for the possibility of indeterminacy and is therefore an im-

provement with respect to Stalnaker's original proposal. Lewis's analysis, however, poses a different problem. The specific indeterminacy of models seems to be difficult to capture in a way that applies universally to all models. Salis and Frigg suggest that the particular way in which the world of a model is incomplete seems to require "a tailor-made cross-world similarity metric" such that "the counterfactual conditional $M \Box \rightarrow C$ has no determinate truth value for all the right C s" (2020: 44).

Williamson (2020) finds the objection unconvincing. He notices that in any conversational context many things that are true are also irrelevant and that a notion of relevance as a standard Gricean conversational implicature could be used to explain the sort of indeterminacy that is characteristic of counterfactual conditionals with irrelevant consequents in models. On this proposal, the world of the model is as complete as any other possible world and the consequents that seem to be indeterminate are determinate yet scientifically irrelevant. Furthermore, a semantic analysis of '*in*' could include a stipulated relevance condition such that C is true in the model if and only if two conditions obtain: i) if M then C holds; and ii) C is relevant to M . Williamson states that the latter, however, "is hardly worth the trouble, since the irrelevant truth is scientifically harmless" (2020).

While this may be the case, the fact remains that the conditional claim "if a gas were composed of point particles exerting no intermolecular forces, then Mont Blanc would be the tallest mountain in Europe" is intuitively neither true nor false. One may have independent reasons to preserve CEM and the completeness of possible worlds, and hence reject the intuition. Or one may recognise that scientific models pose a serious challenge to the completeness of possible worlds in the context of scientific modelling and go for a different analysis that does not satisfy CEM. This would be coherent with the face value practice of modelling and the theoretical principle of naturalism, and it would provide an opportunity for the development of a potentially more fruitful analysis of the sort of indeterminacy involved in models.

The second issue raised by an interpretation of modelling in terms of counterfactual imagination concerns epistemic access. Salis and Frigg (2020: 44) notice that there is no general agreement on the epistemology of counterfactual conditionals. Kment originally held that our ability to gain counterfactual knowledge "needs to be based on rules that permit us to determine which propositions are cotenable with a given antecedent" (2006: 288). Any epistemology of counterfactual conditionals needs to identify these rules. Currently, however, there is no general agreement on what these rules are. In particular, these rules should rely on a previous understanding of the similarity relation between possible worlds, which (as mentioned above) is still insufficiently characterised. These problems are inherited by a counterfactual epistemology of models. The set of C s that are true in a model is different in each case. An epistemology of counterfactual conditionals in models needs to build on a previous understanding of the tailor-made cross-world similarity metric for each case or, as Salis and Frigg tentatively suggest, "perhaps we can identify a series of overarching types of metrics for different types of models" (2020: 44).

The final issue raised by the interpretation of modelling in terms of counterfactual imagination is intersubjective access. Many imaginative activities are solitary and idiosyncratic. This is typically the case in the sort of imaginative activi-

ties involved in dreams and daydreams, and in many cases of counterfactual imagination. Modellers, however, work as members of a scientific community. Their imaginative activities have a social dimension that cannot be explained merely in terms of the ways in which individual modellers think in their own subjective and idiosyncratic ways. Godfrey-Smith himself recognises that model-based science “has sociological and formal features, as well as psychological ones” (2006: 728) and emphasises that he is not interested in providing an account of the psychological mechanisms underlying model-based reasoning. Thus, the analysis of the sort of imagination involved in modelling should build on the social practice of model-based science, and hence on the ways in which scientists think and talk about models as members of specific scientific communities. And while the counterfactual imagination may be compatible with this analysis (once the above problems are solved), a framework that builds merely on this kind of imagination does not have the theoretical resources to explain the social dimension of modelling. This social dimension, as I will argue, is the key feature of a different notion, compatible and yet distinct from the counterfactual imagination. This is the notion of make-believe that I will explore in the next section.

4. Make-believe

Salis and Frigg (2020) argue that make-believe is crucial to theoretical modelling. Walton (1990) originally introduced the notion of make-believe as a social imaginative activity with normative and objective content that is determined by the use of props. Props are ordinary objects that make propositions fictionally true in virtue of a prescription to imagine something. They are material objects that can be perceived and shared by different individuals in a context and thereby provide the physical scaffolding that enables the social, intersubjective dimension of make-believe. Effectively, props afford and constrain the imaginative processes of participants in the make-believe by making manifest the relevant prescriptions to imagine.

What is fictional truth? Naturally, many have spelled out the notion of fictional truth—or fictionality—in terms of fictional worlds. The idea comes from the literature on fiction, where storytelling is often construed as an activity that indicates or creates a fictional world. On this view, Mary Shelley’s act of storytelling selects (among the logical space of possibilities) or generates (through her creative imagination) a world where it is true that Dr Frankenstein creates a hideous, intelligent and articulate creature through an unconventional laboratory experiment. This somewhat natural way of thinking about fictional truth as truth in the world of the story is interpreted in two main ways, literal and non-literal—or imaginative. On the literal interpretation, fictional truth is construed as a variety of truth and being fictionally true is being true in a possible (Lewis 1978) or, perhaps, impossible (Berto 2011; Priest 1997) world. This notion of fictional truth as truth in a world fits well with an analysis of modelling in terms of counterfactual imagination, but it also raises similar problems.

On a second, non-literal interpretation, fictional truth is not a variety of truth but a property of the propositions that are among the prescriptions to imagine in force in a fictional story (Eagle 2007; Currie 1990). This alternative notion of fictional truth, which is Walton’s (1990) preferred notion, is often paraphrased in terms of correctness with respect to the prescriptions to imagine in force in a par-

ticular game of make-believe and has normative and objective features. It is normative because it depends on the rules that guide the imaginings of participants in the game. It is objective because it is independent of the individual imaginings of participants who may or may not conform with the prescriptions to imagine in force in a certain game of make-believe. Furthermore, on Walton's account, fictional truths divide between primary fictional truths and implied fictional truths of the game. Primary fictional truths are the initial assumptions of an episode of make-believe and they are generated directly from the props. Implied fictional truths are inferences generated indirectly from the primary fictional truths via principles of generation (more on these in the next Section).

As stated above, props are ordinary objects that can be perceived and shared by different individuals in a context. What sort of props are involved in a literary work? It is common to indicate, vaguely, the literary work of fiction as the prop. But we can be more specific and say that the concrete tokens constituting the text of a literary fiction are the props that prescribe to imagine in certain ways. These are concrete marks on paper, a computer screen or a tablet, which can be perceived and shared by different individuals in a context. In some cases, they can also be the concrete sounds produced by someone reading a text aloud, hence enabling an audible rather than visual experience of the text. These visible marks (or audible sounds) are the props that enable and constrain the intersubjective and social dimension of make-believe in literary fictions.

These ideas contribute an explanation of model building and model development. Let us start from model building. A scientist builds a model by specifying a model description—the prop—that prescribes certain imaginings. Like the text of a fictional story, the model description, which involves a linguistic and mathematical description, is constituted by concrete, physical marks that can be perceived and shared in a context. These perceptible marks provide the physical scaffolding that make the social dimension of modelling possible. They can be shared by different scientists in a context, hence enabling intersubjective communication within the scientific community and providing tools for the investigation of particular issues.

Similarly to the text of a story, the model description constrains the model's assumptions, or primary fictional truths, coherently with the model's prescriptions to imagine. These prescriptions to imagine involve the attribution of physical properties that only concrete objects can have, yet there are no such objects. For example, the model description of the ideal gas prescribes imagining that the molecules composing the gas are point particles having no volume of their own and bouncing against each other in elastic collisions. In this way, scientists build an imaginary system wherein imaginary gas molecules interact under imaginary conditions. This imaginary system emerges from the propositions that are among the prescriptions to imagine of the model. Hence, it is natural to interpret model building as a cognitive process that is enabled by a use of imagination that diverts from reality in some respects and to certain degrees for the purpose of building a surrogate, imaginary system.

What sort of object is this imaginary system? Realists about model systems argue that they are abstract created entities (Contessa 2007; Giere 1988). Antirealists hold that there are no model systems (Frigg 2010; Salis 2020a). Imaginings have no ontological commitments. So, for example, imagining a witch or telling

a fictional story about some witch do not commit to the existence of any witch.² Similarly, imagining an ideal gas or specifying a linguistic and mathematical description of an ideal gas in the imagination do not commit to the existence of any ideal gas. Walton's theory is compatible with both realism and antirealism about fictional entities. Personally, I have a strong preference for antirealism and I therefore assume that there are no model systems. What follows from this is that model systems are built in the imagination, without any commitment to their existence. Hence, there are no model systems.

Yet, there are models. This much seems undisputable. So, what are they? The term 'model' is often ambiguous between different uses. Sometime it is used to refer to the model system. A realist about model systems can endorse this interpretation of the term 'model' and argue that models are abstract objects. Realism about model systems, however, should be motivated by theoretical considerations that do not depend upon this particular problem. As stated above, I assume antirealism and hold that there are no model systems. This together with the assumption that models are model systems entail the absurd consequence that there are no models. Some other time the term is used to refer to the model description. But a mathematical equation or a string of linguistic symbols are not a model of anything unless they are interpreted in certain ways and according to certain conventions. So, a model description on its own is not a model. However, a model description together with its interpretation (its propositional content) can be identified with the model. On this view, which is the one I favour, a model is akin to a fictional story that the scientist tells by employing certain symbols (linguistic or mathematical) interpreted according to certain conventions.

The propositional content of a model can be analysed according to different accounts depending on one's theoretical stance. Descriptivist accounts will analyse it in terms of general propositions with a uniqueness condition where the description involves apparent reference to a particular (singular) entity, *à la* Russell (1905). Referentialist accounts will analyse it in terms of general propositions and, where certain singular terms such as proper names are involved, singular propositions (realism) or gappy propositions (antirealism), *à la* Braun (2005), or no proposition (antirealism), *à la* Walton (1990: Ch.10). While philosophers of science have well known descriptivist preferences, choosing over one or the other of these options requires independent theoretical reasons that do not hinge on anything specific to the case of models. For this reason, I will not take a stance on this particular issue.

So, on this proposal, a scientist builds a model (intended in this way) by specifying a model description (the prop) together with its interpretation (the primary fictional truths of the model determined by the model's prescriptions to imagine). These, in turn, specify a model system as the object of study, but only within the make-believe. The model is then developed by eliciting what is implicitly true in

² Of course, there were (and in some regions of the world there still are) societies that believed in the existence of witches. These beliefs, however, are rightly rejected in most advanced societies, which find other ways to express their own sexist and misogynistic stances. Fictions can be about real entities. Imaginings, however, do not commit to the existence of the objects they seem to be about. If they are about real entities, they are so in virtue of the existence of these entities.

it—or fictionally true. This requires going beyond the initial assumptions via principles of generation. Specifying what these principles are is no easy feat. I will discuss this problem in the next Section.

5. Constraints on Imagination

Make-believe is a type of imagination that is constrained by the game's prescriptions to imagine and by the principles of generation. In his critical assessment of Salis and Frigg (2020), Williamson (2020) notices that there is no general agreement on the epistemology of make-believe. Understanding our ability to learn through make-believe requires an investigation into the sort of constraints operating on it, including the principles of generation of implicit truths in the model. Salis (2020b) indicates at least three distinct types of such constraints, architectural, context-specific, and epistemic.

Architectural constraints are determined by the cognitive structure of the imagination and operate on all uses of imagination across different contexts. From the contemporary literature in cognitive science emerge two main architectural constraints, mirroring and quarantining.³ Imagination displays mirroring when imaginings carry inferential commitments that are similar to those carried by isomorphic beliefs—that is, beliefs that have the same propositional content. If I believe that it is raining outside, I also believe that the pavement is wet. Similarly, if I imagine that it is raining outside, I also imagine that the pavement is wet. The inferences we make, however, typically depend on background assumptions and on the specific aims and practical interests that direct our reasoning. Thus, mirroring interacts with context-specific constraints to determine the sort of inferences that are allowed in particular episodes of imagination.

Quarantining is displayed when imaginings do not entail beliefs and do not guide action in the real world. In other words, quarantining guarantees that imaginings have effect only within an imagined episode. For example, if I believe that it is raining outside, and I have a desire not to get wet, I will pick up my umbrella on my way out of the house. But if I merely imagine that it is raining outside, I will not act in the same way. This does not mean that nothing of real-world importance can be learned through imagination. Learning about reality through imagination, however, requires exiting the imagination and exporting what one has learned outside of it and into reality. One can study the ideal gas model without automatically learning anything of real-world importance. Gaining knowledge of empirical truths about real world gases requires exporting what one has learned in the imagination onto reality.

While architectural constraints operate on all uses of imagination through different contexts, context-specific constraints are determined by disciplinary conventions and interpretative practices. Individuals who engage in these practices imagine in ways that are specific to the practices themselves. Context-specific constraints correspond to Walton's principles of generation. They are the constraints that enable the generation of implicit truths in a game of make-believe. Inspired

³ Salis and Frigg (2020) identify mirroring and quarantining as two key features of propositional imagination (together with a third one, which is the typical freedom of imagination). See also Leslie (1987), Nichols (2004), and Nichols and Stich (2003) for the original discussion of mirroring and quarantining based on experimental and theoretical research in cognitive psychology and philosophy of mind.

by Lewis (1978), Walton (1990) identifies two main principles of generation, the reality principle and the mutual belief principle.⁴

The reality principle keeps the world of the game as close as possible to the real world. Effectively, this principle relies on the notion of closeness that is key to Lewis's and Stalnaker's standard analyses of counterfactual conditionals. This brings some, although not all, of the aforementioned problems into the framework of make-believe. First, the notion of closeness—reality orientation or similarity—is left unconstrained and mysterious because it is insufficiently characterised. Second, within the framework of make-believe, the reality principle could be implemented without commitment to the completeness of possible worlds. This is because Walton's (1990) appeal to reality orientation is quite loose and does not commit to the completeness of fictional worlds. But as Salis and Frigg (2020) emphasise, different models are incomplete in their own specific ways, which raises the issue of how to provide the right cross-world similarity metric for each particular case. Third, there is no general agreement on the rules that enable us to determine which co-inferences are allowed by a given antecedent. These rules should rely on a previous understanding of the notion of similarity of worlds, which is currently unavailable.

The second principle identified by Walton is the mutual-belief principle, which imports the mutual beliefs of the members of the community in which the game originated. Beliefs are of many different kinds. In the context of modelling, theoretical beliefs and experts' opinions are fundamental for drawing certain inferences within particular models. More context-specific constraints are also possible and new research through historical and contemporary case studies may contribute a better understanding of what they are. Among them are mathematical constraints provided by the particular mathematical tools deployed in a model, interpretations of data, and more.

Finally, epistemic constraints are determined by the particular sort of knowledge we want to acquire. In the context of modelling, there are two main types of knowledge gained through imagination, knowledge of the imaginary scenarios described by model descriptions, and knowledge of empirical truths about reality. These different types of knowledge correspond to two different types of claims generated through imagination in modelling, knowledge claims about the imaginary system specified by the model and knowledge claims about reality.

Knowledge claims about imaginary systems are the claims scientists make within a game of make-believe, such as "the ideal gas is composed of point particles" (in the ideal gas model). These claims are produced within the make-believe, not without it. They are internal claims about the ideal gas (the imaginary system), not external claims about the model (the complex entity constituted by model description and model content). Scientists merely imagine the content of these claims (rather than believing it), which are merely fictionally true (rather than genuinely true).

But what sort of justification do scientists have to make these claims? In the traditional theory of knowledge, justification has the special role of ensuring that "a true belief isn't true merely by accident" (Steup 2018). A belief that p is justified if and only if there are some grounds that properly increase the probability that it

⁴ See also Evans (1982) for a classical discussion of these two principles within the framework of make-believe, and Friend (2016) on the reality principle and a different take on the aboutness of fictional stories.

is true. When we think about the notion of justification in the context of knowledge of imaginary scenarios, the question we need to ask is: What sort of grounds probabilify knowledge claims about imaginary systems? Most plausibly, the relevant sort of grounds must depend on specific modelling practices. Mathematical constraints operate on uses of imagination in all theoretical models. Theoretical grounds may play an important justificatory role in many types of models, including macroeconomic models, models in cognitive neuroscience and models in physics. However, they may play a more limited role in mechanistic models in chronobiology and models in medicine. In these cases, empirical grounds (broadly construed) may play a more relevant justificatory role. More fine-grained distinctions about the specific constraints at work in different modelling practices and even in specific models could be made through case studies.

In the ideal gas model, the principles of generation are quite straightforward and they are provided by the mathematical constraints imposed by the model equation. The model assumes that the volume V of the imaginary gas is proportional to the number of moles n . So, when one doubles n , keeping pressure and temperature constant, V doubles too. In this way one learns about the properties of an imaginary gas. Learning about real gases, however, requires exporting what one has learned about the imaginary system outside of the make-believe and onto reality via the formulation of theoretical hypotheses. These hypotheses are of two kinds, model-world comparisons and direct attributions.

Model-world comparisons are claims that scientists make about the model system and the real system of interest. Often, they are based on a relation of similarity, which is usually interpreted as the sharing of certain properties in certain respects and to certain degrees. So, one can claim that the ideal gas and some real gas have similar behaviours in certain respects. For example, one can claim that when one doubles the number of moles n of an ideal gas and a real gas, keeping pressure and temperature constant, the volume V of the two gases will double too. The ideal gas, however, is only a fiction, a useful construct of the imagination. So, it cannot have the sort of properties that it supposedly shares with real gases. More generally, model systems are constructs of the imagination that do not exist (they are creatures of the imagination that inhabits a model's fictional scenario) and therefore cannot have any of the properties that they supposedly share with their targets.⁵ As a consequence, there cannot be any real similarity between models and reality. But then how can we make sense of the common practice of scientists to compare properties of the model system with properties of real systems?

Answering this question requires that we reconceptualise the notion of similarity in terms of imagined similarity, that is, in terms of the attribution of certain properties to model systems in the imagination, and more specifically within a game of make-believe. According to Walton (1990), games of make-believe can be of two main sorts, authorised and unofficial. A game is authorised when its fictional truths are determined by the model's prescriptions to imagine and the relevant principles of generation. For example, the claim "the ideal gas is composed of point particles" is true in the ideal gas model. A game is unofficial when its fictional truths are determined by some *ad hoc* rules. The claim that "when one doubles the number of moles n of an ideal gas and a real gas, keeping pressure and temperature constant, the volume V of the two gases will double too" is true only in an unofficial game of make-believe constrained by *ad hoc* rules combining

⁵ See Hughes 1997 for a similar concern.

the original prescriptions to imagine of the ideal gas model and new prescriptions to imagine determined by the ways in which real gases behave. Real gases and imaginary gases cannot share any properties, so the claim is literally false. But the same claim is fictionally true when assessed from within the unofficial game of make-believe because they share such properties in the imagination. Knowledge claims generated from model-world comparisons can be assessed only within unofficial games of make-believe and therefore involve epistemic constraints that are similar to those involved in knowledge claims about imaginary systems. Their content is the object of imagination rather than belief. And they can only be fictionally true (or false) when assessed within a game of make-believe (even if unofficial).

Typically, however, scientists build and develop models to learn about reality, to gain some better understanding of it and, possibly some new knowledge. This requires stepping out of the imagination through the formulation of theoretical hypotheses that do not involve any reference to imaginary systems. A scientist can claim that “if one doubles the number of moles of a real gas, keeping pressure and temperature constant, the volume doubles too”. This is a hypothesis that is exclusively about a real system and that can be assessed and even tested for truth. This second sort of hypotheses is enabled by the development of the model in make-believe. But it is exported outside of it in the form of a direct attribution to real systems of the properties attributed to model systems in the make-believe. The knowledge claims generated from direct attributions export what one has learned about the model system into reality, they are exclusively about reality and can be assessed for truth. The attitude one has towards their content is belief and the sort of justification they require is typically provided by empirical evidence.

6. Conclusion

In this paper I advocated the view that scientific modelling crucially relies on imagination of the make-believe variety and that this must be constrained in certain ways to enable knowledge of reality. I described the first overarching taxonomy of types of constraints on imagination in modelling, architectural, context-specific and epistemic. And I identified two main varieties of knowledge generated through modelling, knowledge of the model imaginary system and knowledge of reality. One aspect of the proposal that should be emphasised is that the above taxonomy is open and does not exhaust the many possible specific constraints on uses of imagination in particular modelling practices. New research through case studies is required to specify the different context-specific and epistemic constraints at work in different modelling practices. This, however, will be the aim of future work.

References

- Arlo-Costa, H. 2019, “The Logic of Conditionals”, *The Stanford Encyclopedia of Philosophy*, Zalta, E.N. (ed.), <https://plato.stanford.edu/archives/sum2019/entries/logic-conditionals/>.
- Arcangeli, M. 2018, *Supposition and the Imaginative Realm: A Philosophical Inquiry*, New York: Routledge.

- Berto, F. 2011, "Modal Meinongianism and Fiction: The Best of Three Worlds", *Philosophical Studies*, 152, 3, 313-34.
- Braun, D. 2005, "Empty Names, Fictional Names, Mythical Names", *Noûs*, 39, 4, 596-631.
- Cartwright, N. 2010, "Models: Parables v Fables", in Frigg, R. and Hunter, M.C. (eds.), *Beyond Mimesis and Convention. Representation in Art and Science*, Berlin and New York: Springer, 19-32.
- Contessa, G. 2007, "Scientific Representation, Interpretation, and Surrogate Reasoning", *Philosophy of Science*, 74, 48-68.
- Currie, G. 1990, *The Nature of Fiction*. Cambridge: Cambridge University Press.
- Descartes, R. 1985, *Meditations on First Philosophy*, ed. and trans. J. Cottingham, Cambridge: Cambridge University Press.
- Eagle, A. 2007, "Telling tales", *Proceedings of the Aristotelian Society*, 107, 125-47.
- Evans, G. 1982, *The Varieties of Reference*, Oxford: Oxford University Press.
- Findlay, A. 1948, *A Hundred Years of Chemistry*, 2nd ed., London: Duckworth.
- Friend, S. 2016, "The Real Foundation of Fictional Worlds", *Australasian Journal of Philosophy*, 95, 1, 1-14.
- Frigg, R. 2010, "Models and Fictions", *Synthese*, 172, 251-68.
- Frigg, R. and Hartmann, S. 2018, "Models in Science", in Zalta, E.N (ed.), *The Stanford Encyclopedia of Philosophy*, <<https://plato.stanford.edu/archives/sum2018/entries/models-science/>>.
- Giere, R.N. 1988, *Explaining Science: A Cognitive Approach*, Chicago: The University of Chicago Press.
- Godfrey-Smith, P. 2006, "The Strategy of Model-Based Science", *Biology and Philosophy*, 21, 725-40.
- Godfrey-Smith, P. 2009, "Models and Fictions in Science", *Philosophical Studies*, 143, 101-16.
- Godfrey-Smith, P. 2020, "Models, Fictions, and Conditionals", in Godfrey-Smith, P. and Levy, A. (eds.), *The Scientific Imagination*, Oxford: Oxford University Press, 155-77.
- Harré, R. 1988, "Where Models and Analogies Really Count", *International Studies in the Philosophy of Science*, 2, 118-33.
- Hughes, R.I.G. 1997, "Models and Representation", *Philosophy of Science*, 64, 325-36.
- Kment, B. 2006, "Counterfactuals and the Analysis of Necessity", *Philosophical Perspectives*, 20, *Metaphysics*, 237-302.
- Knuuttila, T. and Voutilainen, A. 2003, "A Parser as an Epistemic Artefact: A Material View on Models", *Philosophy of Science*, 70, 1484-95.
- Leslie, A. 1987, "Pretense and Representation: The Origins of 'Theory of Mind'", *Psychological Review*, 94, 4, 412-26.
- Levy, A. 2015, "Modeling Without Models", *Philosophical Studies*, 172, 3, 781-98.
- Lewis, D.K. 1973, *Counterfactuals*, Oxford: Basil Blackwell.
- Lewis, D.K. 1978, "Truth in Fiction", *American Philosophical Quarterly*, 15, 1, 37-46.
- Lewis, D.K. 1979, "Counterfactual Dependence and Time's Arrow", *Noûs*, 13, 455-76.
- Magnani, L. 2009, *Abductive Cognition*, Berlin and New York: Springer.

- Mäki, U. 1992, "On the Method of Isolation in Economics", *Poznań Studies in the Philosophy of Science and Humanities*, 26, 316-51.
- Mäki, U. 2005, "Models are Experiments, Experiments are Models", *Journal of Economic Methodology*, 12, 303-15.
- Maxwell, J.C. 1965, *The Scientific Papers of James Clerk Maxwell*, Niven, W.D. (ed.), New York: Dover.
- Morgan, M. 1999, "Learning from Models", in Morgan, M. and Morrison, M. (eds.), *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press, 347-88.
- Nersessian, N. 2008, *Creating Scientific Concepts*, Cambridge, MA: The MIT Press.
- Nersessian, N. 2009, "Conceptual Change: Creativity, Cognition, and Culture", in Meheus, J. and Nickles, T. (eds.), *Models of Discovery and Creativity*, Dordrecht: Springer, 127-66.
- Nichols, S. 2004, "Imagining and Believing: The Promise of a Single Code", *Journal of Aesthetics and Art Criticism*, 62, 129-39.
- Nichols, S. 2006, "Just the Imagination: Why Imagining Doesn't Behave like Believing", *Mind and Language*, 21, 4, 459-74.
- Nichols, S. and Stich, S. 2003, *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*, New York: Oxford University Press.
- Norton, J. 1991, "Thought Experiments in Einstein's Work", in Horowitz, T. and Massey, G.J. (eds.), *Thought Experiments in Science and Philosophy*, Lanham (MD): Rowman & Littlefield, 129-48.
- Priest, G. 1997, "Sylvan's Box: A Short Story and Ten Morals", *Notre Dame Journal of Formal Logic*, 38, 4, 573-82.
- Russell, B. 1905, "On Denoting", *Mind*, 14, 479-93.
- Salis, F. 2019, "The New Fiction View of Models", *The British Journal for the Philosophy of Science* (online first), <https://doi.org/10.1093/bjps/axz015>
- Salis, F. 2020a, "Scientific Discovery through Fictionally Modelling Reality", in Ippoliti, E. and Nickles, T. (eds.), *Scientific Discovery and Inferences*, *Topoi*, 39, 927-37.
- Salis, F. 2020b, "Of Predators and Prey: Imagination in Scientific Modeling", in Moser, K. and Sukla, A. (eds.), *Imagination and Art: Explorations in Contemporary Theory*, Leiden: Brill.
- Salis, F. and Frigg, R. 2020, "Capturing the Scientific Imagination", in Godfrey-Smith, P. and Levy, A. (eds.), *The Scientific Imagination*, Oxford: Oxford University Press, 18-50.
- Sorensen, R. 1992, *Thought Experiments*, New York: Oxford University Press.
- Spaulding, S. 2016, "Imagination through Knowledge", in Kind, A. and Kung, P. (eds.), *Knowledge through Imagination*, Oxford: Oxford University Press, 208-26.
- Stalnaker, R. 1968, "A Theory of Conditionals", in Rescher, N. (ed.), "Studies in Logical Theory", *American Philosophical Quarterly*, Monograph Series, Vol. 2, Oxford: Blackwell, 98-112.
- Stalnaker, R. 1981, "A Defence of Conditional Excluded Middle", in Harper, W.L., Stalnaker, R. and Pearce, G. (eds.), *IFS: Conditionals, Belief, Decision, Chance and Time*, Dordrecht: Reidel, 87-105.
- Stalnaker, R. 1986, "Possible Worlds and Situations", *Journal of Philosophical Logic*, 15, 1, 109-23.

- Steup, M. 2018, "Epistemology", in Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/archives/sum2018/entries/epistemology>.
- Suárez, M. 2004, "An Inferential Conception of Scientific Representation", *Philosophy of Science*, Supplement, 71, 767-79.
- Sugden, R. 2009, "Credible Worlds, Capacities and Mechanisms", *Erkenntnis*, 70, 3-27.
- Swoyer, C. 1991, "Structural Representation and Surrogative Reasoning", *Synthese*, 87, 449-508.
- Thomson-Jones, M. 2010, "Missing Systems and the Face Value Practice", *Synthese*, 172, 283-99.
- Toon, A. 2012, *Models as Make-Believe: Imagination, Fiction, and Scientific Representation*, Basingstoke: Palgrave Macmillan.
- Tyndall, J. 1868, *Faraday as a Discoverer*, London: Longmans, Green & Co.
- Walton, K. 1990, *Mimesis as Make-Believe*, Cambridge, MA: Harvard University Press.
- Weisberg, M. 2007, "Who is a Modeler?", *British Journal for the Philosophy of Science*, 58, 2, 207-33.
- Weisberg, M. 2013, *Simulation and Similarity: Using Models to Understand the World*, New York: Oxford University Press.
- Williamson, T. 2005, "Armchair Philosophy, Metaphysical Modality, and Counterfactual Thinking", *Proceedings of the Aristotelian Society*, 105, 1-23.
- Williamson, T. 2020, "Review of The Scientific Imagination, by Arnon Levy and Peter Godfrey-Smith", *Notre Dame Philosophical Reviews*, <https://ndpr.nd.edu/news/the-scientific-imagination-philosophical-and-psychological-perspectives/>