The Role of Models in Modal Modeling

Abstract

Scientists often make and seek to justify modal claims. In this paper, we focus on the role of models in the justification of such claims. In particular, we carve out and distinguish three types of roles that scientific models can play in modal justification: providing concretized situations for compatibility tests, allowing for comparisons between particular systems, and functioning as repositories for the imagination. We illustrate each of these roles with actual modeling practices. We conclude that models are *enablers* of modal justification: they play central roles in the production of scientific modal knowledge, even when the justification strictly speaking comes from various types of empirical background knowledge that informs the modeling practice. Furthermore, examining these roles reveals a plurality of modeling patterns in science, and adds to increasingly pluralist views about model representation.

1. Introduction

Philosophers of science have documented a large number of practices where scientists make (or should be interpreted as making) modal claims with reference to scientific models. Typical examples include *how-possibly explanations*, which includes modeling possible features or causal histories of actual targets (see Verreault-Julien 2019 for an overview); *exploratory* or *hypothetical modeling*, which represents possible features of a putative target lacking widely accepted theoretical descriptions (see Gelfert 2019 for an overview); and *dispositional modeling*, where capacities or susceptibilities are ascribed to real world targets (see Nguyen 2020 for discussion).

We are interested in the epistemology of such *modal modeling*. Recent papers have discussed the conditions under which one may justifiably infer a modal conclusion from scientific modeling (Sjölin Wirling and Grüne-Yanoff, 2021; reference omitted). There, drawing on the modal epistemology literature, three broad justificatory strategies have been identified: (i) coherence with background theory, (ii) relations of relevant similarity, and (iii) imaginability, and argued that existing accounts of modal modeling are usefully seen as engaging one or several of these accounts of how modal claims more generally can be justified.

This raises the question of just what the role of the *models* are, since the modal epistemology literature makes little reference to scientific models. Yet, modeling is central to contemporary science and a lot of modal scientific inferences and practices clearly do involve models. In this paper, we attend to the role of models in the justification of scientific modal claims, since we consider it important to make sense of these practices. We propose three different roles for models in the justification of scientific modal claims, using the broad justificatory strategies (i)-(iii) in carving out these roles for models, and illustrating each with cases from the modal modeling literature. Our ambition is that this will serve to completement both extant modal epistemology accounts, which largely ignore the role of models despite often suggesting that science play an important role in producing modal knowledge, as well as common views on modeling in the philosophy of science, which often center only on models' similarity with actual targets in accounting for their epistemic usefulness.

A caveat before we begin: we assume that a modal empiricist view is central to the understanding of any modal justification in scientific modeling. That is, in modal modeling, scientists will be tapping into *some* kind of empirical background knowledge that, in effect, contributes to the justification. When we distinguish between justificatory strategies is thus rather the *form* of this background knowledge (e.g. propositional/non-propositional, general/particular) that separates them. And as we shall see, they also cast models in different roles – thus implying different modeling practices – when it comes to probing this background knowledge in order to draw modal conclusions.

2. Models as compatibility tests

An intuitive idea is that we come to know what is possible through knowledge of abstract principles that constrain possibility. But knowledge of just what abstract principles? The modal epistemology literature gives at least three different (but not mutually exclusive) answers. First, a popular view in recent years is that knowledge of what is possible depends on having access to *essentialist, or constitutive knowledge* (e.g. Lowe 2012; Jago 2018; Mallozzi 2018). Translated to the context of modeling, one is justified in taking a model to represent a possible object/system if the model can be shown to satisfy the essentialists requirements for the object/system's kind. Second, it is natural to think that *knowledge of laws* (e.g. laws of nature or "metaphysical laws") have an important role to play (e.g. Kment 2014; 2021).[1] Translated to the context of modeling, one is justified in taking a model to represent a possibility if one can show that it is compatible with all laws that apply to it. Third, some have argued that possibility knowledge is downstream from justified *theories* about the relevant phenomena or entities. Importantly, while the two previous accounts tend to talk of knowing a possibility claim to be *compatible* with the constraint, theory-based accounts often

[1] For a different take on the intimate relation between modal knowledge and knowledge of laws (of nature), see Wilson (2020).

envisage a stronger tie, in terms of e.g., derivation or implication (Bueno and Shalkowski 2014; Fisher 2017).

This broad type of justificatory strategy can be fairly straightforwardly transferred to a scientific context. In such contexts, any constitutive or nomological principles would be part of a theory, so in what follows we will concentrate on theory-based accounts. The philosophy of science literature has a lot to say of the relation between theory and model. Models are used to probe or test theory (Redhead 1980), they serve as "nomological machines" that conceptualize phenomena in ways that make theoretical principles applicable to them (Cartwright 1997), and they mediate between theories and the states of the world that the theory applies to (Morgan and Morrison 1999). Crucially, in each of these accounts, models complement theories by filling in details of particular situations that remain incompletely specified in the theory – even though model systems might not be particulars themselves, they "would be concrete if real" (Godfrey Smith 2009, 104). None of these accounts address models as justifiers for modal claims, but they can, we think, easily be so extended in order to illustrate what role models play given a background theory-based account of modal justification.

In particular, we suggest, this first role for models is as a tool for testing whether the theory and its implications are compatible with some scenario in which a prospective possibility *p* is true, or more strongly, whether *p* is derivable from the theory. The model provides a particular context in which *p* is true, and where relevant theoretical principles are implemented together. The result of this test then provides the justification for the possibility of *p*. Nevertheless, the model itself is not the source of the justification – that is instead the background theory. Using a model, one can certainly test whether *p* is possible *according to a theory* one knows to be false – but that test doesn't justify the claim that *p* is possible. The justification for the possibility claim piggy-backs on the justification of the background theory.

This is not to say that the model is dispensable or unimportant. Quite to the contrary, it functions as an important type of mediator in which the modal implications of existing background knowledge are probed and discovered. Perhaps a mere imaginative exercise might suffice in some situations to explore and test that the compatibility of some *p*-scenario with certain theoretical principles and laws – some of the modal epistemology literature seems to suggest as much (e.g. Fischer 2017, 26-27, see also Nichols 2006). The rigor of such practices is limited, however, by the number of interacting factors that humans can mentally represent at the same time. Furthermore, it is difficult to make such imaginative exercises intersubjectively accessible to other researchers. Modeling practices extend these simultaneous representation abilities and allow for intersubjective accessibility. The use of models thus increases rigor and intersubjective accessibility of such compatibility tests.

A good illustration of this mediating role in the context of modal modeling can be found in Michela Massimi's (2019) work on exploratory modeling practices, where she argues that the epistemic import of some such practices consist in delivering knowledge of what is possible. According to Massimi's analysis, they do so by way of what she calls *physical conceivability*. Physical conceiving is a mode of imagining – imagining constrained by knowledge of *nomological principles*, in particular.

> *p* is physically conceivable for an epistemic subject S (or an epistemic community C) if S's (or C's) imagining that *p* not only complies with the state of knowledge and conceptual resources of S (or C) but it *is also consistent with the laws of nature known by S (or C)* (2019, 872, our emphasis).

The purpose of (some) exploratory models is thus to aid scientists in physically conceiving some hypothetical *p*. Massimi illustrates this with an example where modeling hypothetical theoretical postulates using a technique called the pMSSM-19 is used in the search for supersymmetric (SUSY) particles. Particle physics has theorized a hypothetical entity known as the SUSY particle, in order to account for certain gaps in the Standard Model. However, scientists have not yet been able to confirm whether there actually are any SUSYs in nature, and they do not know what properties they would have if they did exist. Exploratory modeling ultimately helps in this endeavor by outlining different ways in which is physically possible that a SUSY particle exists.

Basically, the idea is that scientists are justified in regarding models that can be conceived under the constraint of relevant nomological principles (R parity conservation and consistent electroweak symmetry breaking, among others) as representing objectively possible ways for a SUSY to be, if it existed. This information – a set of possibilities – is then used to design experiments that can help further whittle down the set in the search for actual SUSYs. What is important for our purposes is that the inference in this instance of modal modeling runs from the theory (including a set of laws), through their implementation or compatibility in an imagined concrete scenario represented by the model, to the conclusion that the models passing the test represent possibilities. That is, the model's primary function is to allow for the concretization & joined implementation of the abstract principles in which the empirical background knowledge is contained. It is only in the model that this test can be rigorously performed in an intersubjectively accessible way, because the model is the place where abstract principles and bits of information are concretized into a situation in which the relevant *p* is true. This shows that models play a crucial role in the modal inference despite the fact that the justificatory work as such is done by the theoretical principles that constrain the model.

3. Models as particular individuals allowing for comparison

According to another, recently quite influential idea in the modal epistemology literature, one can draw justified conclusions about what is possible for some target *t* on the basis of knowledge of what is actually the case with other, distinct entities that are relevantly similar to *t*. For instance, one can know that it is possible for this table to break because one has seen other tables, relevantly similar to this table, actually break. This role of similarity-judgements as supporting modal reasoning has been highlighted by Roca-Royes (2017) in particular.[2] Obviously, one's knowledge about entities similar to *t* may be encoded in a theory that one accepts, or one may have abstracted knowledge of laws from one's empirical knowledge, but it is important to note that this is not necessary for similarity-based justification. An important point for Roca-Royes' is that similarity-based modal epistemology is not supposed to be dependent on knowledge of essentialist principles, for instance (Roca-Royes 2017, 223). All that is required for grounding inferences about possibilities for the target individual is experiential knowledge of the properties of *particular individuals* and an ability to compare them.

That similarity-judgements should also play a role in supporting modal conclusions when scientific models are involved is not very surprising. Appeal to similarity in some form or other as a justificatory basis for model-target inferences is a venerable tradition in the philosophy of science. The basic idea here is that one can learn about target *t* by studying model *m* in virtue of the fact that *t* is similar to *m* in relevant respects. The relevant similarity may be e.g. partial, structural, with respect to properties, behavior, or causal mechanisms (Giere 2010; Godfrey-Smith 2009; M. Weisberg 2013). What similarities *are* relevant in a given modeling context is plausibly determined to a large part by the interests and purposes that modelers have in mind. Moreover, while scientific theory and knowledge of laws clearly inform the relevant similarity-judgements when available, this is not necessary – just as in the case of similarity-based modal epistemology. Reasoning from knowledge about the model to conclusions about the target can proceed independently of and in the absence of accepted scientific theory about the entities in question, on basis of knowledge about the model's and the target's respective properties. The role of the model here is to act as an *epistemic surrogate* for the target(s). Instead of directly studying and theorizing about some target, a modeler

[2] See also (Dohrn 2019; Leon 2017).

makes the strategic decision to instead construct a distinct, typically simpler and idealized, system that "stands in" for the target of interest, and tries to understand *its* workings first, and *then* one draws conclusions about the target.

The similarity-account of how scientific models relate to their targets is not tied to *modal* modeling, but it too can quite naturally be so extended. Indeed, it points us to a second role for scientific models to play in modal reasoning, which dovetails neatly with Roca-Royes' similarity-based modal epistemology. Simply put, the model functions as a "concrete" context that allows for study, manipulation and *comparison* with the distinct particular target individual – just as in modeling more generally, according to the similarity-account, but in the cases at hand the conclusions drawn about the target are modal conclusions.

Such comparisons might work as follows (Sjölin Wirling forthcoming). If we want to find out whether some target system of interest T could possibly be F (perhaps despite being actually not-F), we could in principle investigate other actual systems that are F, compare them to T and see whether it and any of these F-systems are relvantly similar (or, we could investigate systems known to be relevantly similar to T and check whether any of them are F). But if there are no F-systems relevantly similar to T, we can still find out whether T could possibly be F by constructing a model system M, attempt to realize F in M, and then compare M to T in order to see whether they are relvantly similar. If they are, we seem to have reason for thinking that T could possible be F. Here, the role of the model is to replace the actual, relevantly similar individuals on which we could have drawn on if they were available. Again, the model is a particular individual which allows study, manipulation and *comparison* with the distinct target individual, and this grounds modal conclusions about the target.

To clarify, the model contributes to the justification of modal claims based on similarity-judgements by providing the basis for the comparison with the target. Thus, the model will often be indispensable in the sense that scientists often turn to models not just for convenience but because there are no actual systems known to be F, or because for some reason or other (e.g. practical, economical, ethical, or as a matter of principle) scientists cannot study actual potential F-systems in order to compare them to the target. Models enable the *extension* of an intuitive form of reasoning from actual individual to modal conclusions about other actual individuals, to cases where no actual (or no actual and relevantly similar) individuals are known to realize the property of interest. They are constructed "epistemic counterparts" (cf. Roca-Royes 2017, pp. 226) of the target systems we are interested in.

It may now be objected that even if there is nothing in principle wrong with the picture just sketched, it does not fit well with many examples of modal modeling discussed in the literature. Often, modal (re-)interpretations of modeling practices are prompted by an apparent *lack* of model-target similarity. From the perspective of a similarity-account of modeling, this makes them unsuited to support conclusions about what actually is the case with some target system, and so their epistemic relevance must be otherwise accounted for. However, although it is true that much of the recent interest in modal modeling is driven by cases where the similarity-based account is said to fall short, this does not prevent there from being other cases of modal modeling where similarity is arguably doing the work. These may become more visible once the phenomenon of modal modeling is recognised and we zoom out from the class of models that originally drew attention to this phenomenon, and look to other modeling practices that may also involve modality.

Moreover, we think that models play this role of allowing for informative comparisons between particulars that ground justified modal conclusions even in some cases from the existing modal modeling literature, even if the practice in which this is embedded looks very different from a more standard case of similarity-based possibility reasoning based on modeling. We will illustrate this with Batterman and Rice's discussion of "minimal models". Minimal models are simple, highly idealised, abstract, mathematical models that differ in a number of respects from any actual target system, including what are known to be core causal mechanisms and processes. Yet these models apparently are important to our understanding of various actual, often complex and messy, systems. In part, Batterman and Rice's analysis of minimal models is motivated by the alleged failure of standard similarity-accounts to explain the epistemic import of these minimal models in terms of shared features between model and target.

One example of a minimal model is R.A. Fisher's (1930) sex ratio model in evolutionary biology. It addresses a phenomenon that has interested many evolutionary biologists, namely why so very many populations, otherwise extremely dissimilar from one another, all have a sex ratio of approximately 1:1 (i.e. there is a roughly equal number of males and emales in the population). Fisher's model is a dynamic mathematical model which considers three generations of a population, constrained by a number of assumptions. The gist of it can be captured as follows. If the population diverts from the 1:1 sex ratio, so that there is an overweight of, say, female offspring, parents that produce more offspring of the minority type (i.e. male), will have a fitness advantage – because their male offspring will have more mating opportunities. This will lead to an increase of male births in the population since the strategy of producing male offspring is fitter. However, as the sex ratio starts to approach 1:1, this fitness advantage will diminish and finally disappear. Then there is no motivation to continue the overproduction of male offspring. Since the same thing happens when the roles are reversed (i.e. when one starts with an overweight of males in the population), the 1:1 sex ratio is the stable equilibrium towards which natural selection will see that evolving populations end – under the assumptions constraining the model, that is. But many of these assumptions do not hold true of actual populations – nevertheless, the Fisher model help explain the sex ratio in a dazzlingly wide range of actual populations.

According to Rice (2018, 2019), a minimal model like Fisher's is best seen as an instance of modal modeling. In particular, such models deliver *counterfactual knowledge* of actual target systems. In order to see how, we need to consider the Fisher model as part of a larger process which they call *delimination of a universality class*. A universality class is defined in terms of (disposition to exhibit) a certain macrobehavior. This is described mathematically as a disposition to flow towards a given fixed point in a state space under certain transformations. The process of delimiting a universality class starts with the scientists' considering a "space of possible systems", which are distinguished by their constitutens, their degree of simiplicity and the idealizations involved, assumptions about the internal dynamics of the system, and so on. Some of the systems making up this possibility space will be model systems like Fisher's model and variations thereof, whereas others will be (representations of) actual populations. Each of these will exhibit some macrobehavior, and some of them will exhibit the same. The scientists may then start to probe this possbility space in different ways, using various mathematical operations. In essence what is at issue is a form of robustness analysis: to check across what variations in assumptions, constituents, etc. that the feature, function, or behavior of interest – in this case a 1:1 sex ratio – is maintained. Systems that continue to exhibit the same macrobehavior are identified as belonging to the same universality class.

Now, what scientists discover when they succeed in delineating a universality class is that a number of very diverse systems flow towards the same fixed point (i.e. are disposed to exhibit the same macrobehavior). They can then check for common features to the systems in the universality class, and in the case at hand find that they all have what economists call a linear distribution cost – in other words, for all systems in the universality class of systems tending towards a 1:1 sex ratio, sons and daughters cost an equal amount of resources to produce. Moreover they can also, by further probing the space of possible systems, see that absent this feature, the sex ratio will not tend towards a 1:1 equilibrium. The upshot of all this is, according to Batterman and Rice, that scientists have gained knowledge in the form of counterfactual conditionals pertaining to various actual populations.[3] For instance, they can conclude that if, say, a deer population Q had a very different size than it actually does, Q would still have a 1:1 sex ratio; that if phenotypes were inherited in a very different way in a given shark population R, R would still have a 1:1 sex ratio; that if sons cost much more resources than daughters to produce in a given lynx population S, S would not have had a 1:1 sex ratio; and so on. Indeed, if the probing of the possibility space is thorough enough, scientists can claim to know that linear distribution cost is necessary for a 1:1 sex ratio (Batterman and Rice 2014, 361).

The process described above is, of course, quite different from similarity-based reasoning from model to (actual) target: scientists are not constructing a model system meant to be relevantly similar to e.g. Q, study it and then compare the two systems in order to draw conclusions about Q. If anything, it is rather that we *find out* through this modeling exercise that Q and the Fisher model *are* relevantly similar – despite appearances – and we don't know that prior to the mathematical operations through which the universality class was delineated. Nevertheless, here too comparisons between particulars are central to justifying the relevant modal claims, and the *the role of the model* is again to allow for such comparisons. Another difference is that the comparison is not directly and exclusively between a model and a target, but between the large number of systems in the pre-defined possibility space – still, the role of models that make up the possibility space is to allow for comparisions between the systems in the search to delimit the class of models that are relevantly similar, i.e. exhibit the same macrobehavior.

Next, once one has on hand the universality class, it is through assessing models and actual targets for similarities and dissimilarities, in light of the fact of their disposition (or lack thereof) to display a certain macrobehavior that the modal knowledge pertaining to actual systems is acquired. That is how scientists can find out what the necessary condition(s) for the macrobehavior is – linear distribution cost – and based on that insight they can establish counterfactual conditionals pertaining to actual systems, depending on whether or not the systems are relevantly similar to the

[3] Such counterfactual information is allegedly epistemically valuable because knowledge of counterfactual conditionals is central to scientific understanding (Woodward 2005).

model in this respect.[4] All the way through, a key role of the model(s) is to enable comparison between particulars.[5]

In sum then, although the modeling practices and context are very different between these two instances of modal reasoning on the basis of models, the core role of the model is at heart the same: its function is to enable particular-to-particular comparisons, and assessment of the similarities (and dissimilarities) between them brings out the modal conclusions.

4. Models as repositories for the imagination

An old thought in the epistemology of modality is that modal knowledge is acquired through the imagination. There is currently no agreed upon definition of imagination, and most authors stress the diversity of different kinds of imagination; but most agree that the act of imagining involves representing "without aiming at things as they actually, presently, and subjectively are" (Liao and Gendler 2020). One prominent idea in modal epistemology is that the ability to imagine a state *p* is considered evidence for the possibility of *p* (e.g. Kung 2010; Yablo 1993), another is that the imagination is central to the evaluation of counterfactual claims of the form 'If *a* were the case, then *b* would be the case' (Williamson 2007, see also Byrne 2005). However, close consideration of the imagination suggests that there are several hurdles to the idea that the imagination constitutes "a reliable way of forming a true belief" (Williamson 2013, 119).

First, imagination is considered to be *too liberal*. As the official policy of many patent offices regarding the handling of perpetuum mobile submissions attests (e.g. Section 4.05 of the UKPO Manual of Patent Practice), some people not only can imagine the impossible, but seem to enthusiastically take it as the justification for actionable beliefs. If such imaginative excesses cannot be reined in, then imagination would not be a reliable justification of beliefs in general and modal beliefs in particular. A second, diametrically opposed worry is that imagination is *too constrained*. When identifying possibilities, people often seem to be too tightly constrained by their knowledge of reality. Such "imaginative resistance" (Gendler 2000) might prevent people from creatively

[4] It is true that the relevant disposition, in this case the linear distribution cost, can be realized in very different ways (e.g. by very different features or consituents) across the different systems, but it is nevertheless a similarity, and it is because we judge this similarity to obtain that we can draw the modal conclusions in question.

[5] This is not to say that Batterman and Rice's account collapses into what they call a "feature sharing account", as Lange (2015) argues that it does – see McKenna (2021) for an attempt to rebut that criticism. We are focusing here not on the question of explanation, but on the claim advanced by Rice in subsequent work that the upshot of minimal modeling is counterfactual knowledge. However, there is a clear sense in which model-target similarity is important to the question of in virtue of what e.g. Fisher's model can help inform us about actual target systems (see [reference omitted] for discussion).

identifying relevant possibilities. What is more, people might be subject to systematic biases which possibilities they can imagine. For example, people are more likely to change the last event in a causal sequence, rather than an earlier one; and they more likely imagine them *not* having done something they did, rather than imagine having done something they didn't (Byrne 2005). A third worry, overlapping with the previous two, is that imagination by its nature is disconnected from reality, and thus cannot justify beliefs about reality.[6] Some authors have argued that while modeling imaginary situations might justify possibility claims about these imaginary model systems, a justification of possibility claims about the world "requires exiting the imagination and exporting what one has learned outside of it and into reality" (Salis 2020, 74). Insofar as this exporting cannot be performed by the imagination exercise itself, the justification requires at least something in addition to pure imagination. In sum, while it is widely recognized that the imagination is often useful in modal reasoning it remains a point of serious contention whether it really does any justificatory work on its own, or whether that justification is always downstream from some appropriately justified constraints.[7]

In the philosophy of science, it is agreed that imagination plays *some* important roles in science (Holton & Holton 1978, Stuart 2020) but there is considerable disagreement just what functions it can fulfill. Few would deny that the creation of new ideas, procedures or tools requires imagination. But while some authors do claim that the imagination play a part in epistemological justification also in science, others have strongly resisted such epistemological justification claims, both historically (for an overview, see Daston 1998) and today (French 2020). Deniers typically appeal to the same limitations of the imagination as reviewed in the previous paragraph. Defenders of imagination's justificatory function generally admit that these criticisms apply to *some* imagination exercises. But they argue that for purposes of scientific justification, the imagination can be appropriately constrained so as not to suffer from these shortcomings. Such constraint proposals concern both the starting point of the imagination (its premises or initial conditions) as well as the process of the imagination itself. In some instances of modal modeling, one broad kind of constraint on the imagination are external constraints, for example that the initial conditions and processes of imagination must satisfy relevant theoretical knowledge in the domain of interest, as in Massimi's account of exploratory SUSY modeling described in section 2.1 above (see also Sugden 2000), or that they be sufficiently similar to specific instances of reality. These constraints might guard against the "too wide" worry, as imaginings incompatible with theoretical knowledge

[6] This criticism requires especially subtle formulation with regards to possibility claims, as at least some of these concern non-actual features of the world

[7] For this type of criticism see e.g. (Roca-Royes 2011a; 2011b; Tahko 2012; Fischer 2017, chapter 6).

or certain features of reality would be excluded by them. Furthermore, they might also lend support for the exporting from imagination to reality: because the imagined possibilities are now certified by constraint informed by reality, they can be legitimately be applied to it. Finally, theoretical constraints allow for the construction of possibility spaces: they provide the conceptual tools and the constraints within which all possibilities can be in principle mapped. This allows systematically exploring the possibility space, and thus guards against the "too narrow" worry. These arguments are further supported by empirical evidence that human exercises of imagination tend to respond appropriately to the structure of the real world (D. Weisberg 2020).

However, it seems difficult to differentiate the function of models in theoretically constrained imagination that justify modal claims from those functions discussed earlier in this paper; theoretical or similarity constraints seem to do the justificatory work, and the models serve as (imagined) particular situation where the theoretical assumptions are tested together, or an (imagined) particular individual which can be compared to some target or other model system.[8] While we want to stay neutral on the question of whether the imagination really is a separate source of modal justification, we do think that there is a third type of role for models, and that this role can be brought out by considering a slightly different approach to modal justification by imagination.

Instead of external constraints to be imposed on the imagination, some authors seek to overcome the hurdles of imagination's limitations by focusing on the right *kind* of imagination. In modal epistemology Williamson (2013), for example, distinguishes between *voluntary* and *involuntary* exercises of the imagination, arguing that reliable forms of imagination imagine the antecedent of the conditional voluntary, but the consequent involuntary: "Left to itself, the imagination develops in a reality-oriented way, by default" (Williamson 2013, 118). Kung (2010) distinguishes between *qualitative content* of imagination, which is what we phenomenologically experience (e.g. "see") in the imagining that we conjure up, and *stipulative content* which is a form of intentional "labelling" of the qualitative content, and argues that only the former justifies modal conclusions. A somewhat similar account in the philosophy of science literature is due to Nersessian (1993; 2010). She connects modeling to thought experimenting, a form of imagination, and argues that models serve as means by which thought experimenters represents their scenarios. The imagining agent draws the elements of the scenario from various forms of non-propositional (e.g. sensory) knowledge and

[8] Massimi (2019) distinguishes between law-based and law-guided conceivability. But this distinction is based in differences in modeling practice; she does not address whether there are any differences in justification.

seeks to create a coherent whole. The role of the model is to record this scenario and provide the means by which the imagining agent can "see" the scenario and evaluate its coherence.

To emphasize, this role for the model would be different from the two roles discussed in previous sections. Models are now not the carriers of the constraints imposed on imagination – there are no externally certified rules of inference or principles of generation, represented by the model, which the model user is invited to imagine. Rather the other way around: the model is *the repository of the various steps in the imagination process*. It records results of the imaginings, and thus allows model users to behold the entirety of their imagination exercise, allowing for narrative coherence checks and also providing a basis for further "involuntary" or "qualitative content" imaginings.[9] The generation of the results, recorded in the model representation, are thus driven by the imaginative process itself, not by the model. Let us illustrate this with two examples: prototype modeling in engineering design, and game theoretic modeling in the social sciences.

Engineers in early stages of the design process often build prototypes: physical models of a design idea that has not yet reached production stage. The function of such prototypes, it has been suggested, include reducing ambiguities in design concepts (Lemons et al. 2010), identifying drawbacks of undesirable features present in design ideas (Viswanathan et al. 2012) and detect errors in designers' own internal representations (Ramduny-Ellis et al. 2008). These claims portray the material model as helping to complement, adjust and improve upon a mental representation of the intended design. The designers first imagine a design possibility, develop it in their minds, and then realize this possibility by building a prototype. Building this prototype forces designers to specify their design idea further, thus eliminating ambiguities. Once the prototype is built, it also serves as a comparison to the mental image, allowing designers to test them for coherence. Finally, now that the prototype represents what they have imagined so far (freeing up designers' cognitive capacities), designers can now proceed to the next step, imagining *about* the prototype what might go wrong – for example, whether critical loads have been neglected, connections between parts not included, unnecessarily expensive materials or non-standard parts selected, production time and budget not adequately planned (Viswanathan et al. 2012). These imaginings about the prototype lead to the discovery of new possibilities of errors or needed improvements, which in turn induces changes to the prototype. In iterative processes of imagination and prototype-building

[9] These views stress the materiality and tool-likeliness of models, and are therefore sympathetic to Knuuttila's (2017) artifactual approach to models: "The external representational means in their various modes and media embody model systems thereby extending scientific imagination and reasoning into the artifactual realm."

(Viswanathan et al. 2014 report between 17 and 26 iterations), these error possibilities are explored until a final design is reached.

Of course, at least some of these errors might be detected by applying theoretical design principles to the model, and testing the prototype for relevant similarities to the real target domain. Yet a number of empirical studies show that engineers often do not rely on such strategies, instead using the model for error detection without such formal tests. A recent study of how experts and students perform design tasks shows that a substantial part of prototype changes (30% for experts and 40% for students) were the result of "unarticulated tests" – i.e. interactions with the model that were not originally planned by the team, but rather were spontaneously motivated by considering the prototype itself. The study concludes that "building and testing prototypes helps to supplement designers' incomplete mental models leading them to better ideas" (Viswanathan et al. 2014, 1). At least some of these inferences about possible errors, we conclude, are justified with reference to the designers' imagination exercises. Whether or not the imagination is the basic source of the justification here, what is important is that a distinct role for the model in modal justification is outlined here: the prototype model enables these imaginings, first by representing them, but secondly by inspiring further imagination in interaction with this representation.

The second case is from game theoretic modeling in the social sciences. Game theory is a form of mathematical modeling that has been used in many domains, but predominantly for the purpose of modeling strategic interaction between rational agents. As the notion of rationality is not entirely agreed upon, the theory provides a menu of solution concepts that can be employed to analyze a model of a particular strategic situation (a "game"). There has been an extended discussion of what these models are models of, with many arguing that the models represent exemplary narratives, "in search of observations" (Morgan 2007, Sugden 20011. This implies that an analysis of the stock games in the literature – think "Prisoners' Dilemma", "Chicken", "Tit-for-Tat"- relies on the choice of solution concepts *before* their similarity is compared to any real-world situation. The theory itself does not specify the precise condition of their applicability, hence there must be some extra-theoretical reasons to choose between them. Here, the exemplary narrative plays an important role. It provides a plausible but imaginary back-story that satisfies the properties of the formal game-model (specifying the number of players and strategies, payoff functions, sequence of moves) but also rudimentarily portrays the players' epistemic and motivational context. Consequently, the model narrative gives reasons for choosing a solution concept for the specific imaginary situation represented by the model (Grüne-Yanoff & Schweinzer 2008). The modeler thus employs game theory to first represent a possible situation with the formal tools of the theory. They then use that representation to justify the choice of a solution concept, which in turn is used to analyze the possible outcomes of this imaginary situation. Connections to real phenomena (i.e. whether a specific phenomenon instantiates such a possible situation and thus might result in the analyzed possible outcomes) typically are drawn *after* the imaginary situation has been thus analyzed.

Both cases, we believe, illustrate the use of models as repositories of steps in the imagination process. Engineering prototypes visualize and make tangible design ideas, thus allowing further imaginative steps ("unarticulated tests") to detect possible errors and improvement possibilities, which lead to novel prototypes. Game models represent features of imaginary strategic narratives. These possible narratives are then used to justify the choice of analysis tools for these game models, which in effect become formalized representations of how the modeler imagines a rational agent to choose in such a situation. In both cases, theoretical principles or similarity comparisons to real particulars cannot account for all the inferences and decisions made in the modeling process. They therefore illustrate how models contribute to the justification of modal claims by substantially supporting imagination exercises.

5. Conclusions

We have carved out three distinct roles for scientific models to play in the justification of modal claims and illustrated them with examples from the philosophy of science literature. First, models can function as concretized situations in which the compatibility between abstract theoretical principles and particular states are tested, in order to draw conclusions about the possibility of said states. Second, models can function as particulars that allow for comparisons with target systems, in order to draw modal conclusions about those targets. Third, models can function as repositories for the imagination that allow testing the coherence of imaginings, in support of conclusions about nonactual possibilities that in turn serve as the basis for further imagination exercises.

The need to carve out roles for scientific models in modal justification arose partly from the fact that existing accounts of modal modeling at bottom appear to boil down to one of or a mix between accounts of modal justification more generally, where models have no obvious place. That might be taken to suggest that models are somehow dispensable to modal inferences in science. That is, a worry arose from the mismatch between the centrality of modeling to scientific inferences on the one hand, and the reliance on justificatory strategies that did not include models on the other. In this paper we have endeavored to curb this worry by arguing that in fact models play very central roles in the production of scientific modal knowledge, even when the justification strictly speaking come from various types of empirical background knowledge that informs the modeling practice. Models are enablers of modal justifications.

This enabling function does not give rise to a single use-pattern, though. Even within the domain of modal modeling, we see widely differing practices. Compatibility tests require the construction of a single "concrete" situation from theoretical principles in which the possibility claim is true. Similarity judgments (for modal claims) require the simultaneous construction of (often multiple) epistemic counterparts, so as to allow for comparisons between them and/or with a target. Imagination exercises, in contrast, require the sequential interaction of imaginings and their presentation, leading to series of modified models, as witnessed in the prototype example. Our study of modal justification with models thus provides an explanatory framework for the observed plurality of modal modeling practices.

Furthermore, our study provides novel perspectives on representational function of models. A growing literature has rejected the idea that similarity between model and target should be the general criterion for the representational quality of a model (as defended in M. Weisberg 2013). Instead, the justificatory strategies we identified here either make no use of similarity comparisons at all; or, when they do, the similarity comparison need not be between model and (actual) target, but can be between multiple model systems. Our study therefore adds to increasingly pluralist views about model representation

Finally, let us stress that we do not take this list of three roles to be exhaustive – there are surely other roles that models can play in modal justification, and we encourage further investigation of those. Similarly, we have carved these three roles out against the backdrop of three different broad justificatory strategies, and while we see a good fit between these roles and strategies, we are open to the possibility that there might be different roles for models within each strategy and that the same role may be played within different strategies. In general, much research remains open with respect to the phenomenon of modal modeling, and our hope is that this paper both contributes some insight into how models contribute to modal inferences in science and stimulates further probing by others into this and related questions.

References

Batterman, R. and Rice, C. (2014). Minimal Model Explanation. *Philosophy of Science, 81*(3), 349-376.

Bueno, O. and Shalkowski, S. (2014). Modalism and theoretical virtues: toward an epistemology of modality. *Philosophical Studies, 172*(3), 671-689.

Byrne, R. M. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, Mass.: MIT press.

Cartwright, N. (1997). Models: The blueprints for laws. *Philosophy of Science*, *64*, S292-S303.

Daston, L. (1998). Fear and Loathing of the Imagination in Science. *Daedalus*, *127*(1), 73-95.

French, S. (2020). Imagination in Scientific Practice. *European Journal for Philosophy of Science*, *10*, 1-19.

Dohrn, D. (2019). Modal Epistemology Made Concrete. *Philosophical Studies, 176*(9), 2455-2475.

Fischer, B. (2017). *Modal Justification via Theories*. Cham: Springer International.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

Gendler, T. S. (2000). The puzzle of imaginative resistance. *The Journal of Philosophy*, *97*(2), 55-81.

Giere, R. N. (2010). An agent-based conception of models and scientific representation. *Synthese, 172*, 269-281.

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

Grüne-Yanoff, T. and Schweinzer, P. (2008). The roles of stories in applying game theory. *Journal of Economic Methodology* 15 (2), 131-146.

Holton, G., & Holton, W. (1978). *The scientific imagination: Case studies*. Cambridge: Cambridge University Press.

Kment, B. (2014). *Modality and Explanatory Reasoning*. Oxford: Oxford University Press.

Kment, B. (2021). Essence and modal knowledge. *Synthese* 198 (S8), S1957–S1979.

Knuuttila, T. (2017). Imagination extended and embedded: artifactual versus fictional accounts of models. *Synthese*, https://doi.org/10.1007/s11229-017-1545-2.

Kung, P. (2010). Imagining as a guide to possibility. *Philosophy and Phenomenological Research*, 81 (3), 620-663.

Lange, M. (2015). On 'Minimal model explanations': A reply to Batterman and Rice. *Philosophy of Science*, *82,* 292-305.

Leon, F. (2017). From Modal Skepticism to Modal Empiricism. In B. Fischer & F. Leon (Eds.), *Modal Epistemology After Rationalism* (pp. 247-261). Dordrecht: Springer.

Liao, S. and Gendler, T. (2020). Imagination. In E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition) URL = <https://plato.stanford.edu/archives/sum2020/entries/imagination/>.

Lemons, G., Carberry, A., Swan, C., Jarvin, L., and Rogers, C. (2010). The Benefits of Model Building in Teaching Engineering Design. *Design Studies* 31 (3), 288-309.

Massimi, M. (2019). Two kinds of exploratory models. *Philosophy of Science, 86*(5), 869-881.

McKenna, T. (2021). Lange on minimal model explanations: A defence of Batterman and Rice. *Philosophy of Science,* 88 (4), 731-741.

Morgan, M. (2007). The Curious Case of the Prisoner's Dilemma: Model Situation? Exemplary Narrative? In A. Creager, E. Lunbeck, and M. Norton Wise (eds.) *Science Without Laws* (pp. 157-185) Durham: Duke University Press.

Morgan, M. S., & Morrison, M. (1999). *Models as mediators* (p. 347). Cambridge: Cambridge University Press.

Nguyen, J. (2020). It's Not a Game: Accurate Representation with Toy Models. *British Journal for the Philosophy of Science*, *71* (3), 1013-1041.

Nichols, S. (2006). Imaginative Blocks and Impossibility: An Essay in Modal Psychology. In Nichols, S. (ed.) *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction* (pp. 237-256). Oxford: Oxford University Press.

Ramduny-Ellis, D., Hare, J., Dix, A., and Gill, S. (2008). Exploring Physicality in the Design Process. In *Proceedings of Design Research Society Conference*, Sheffield, UK.

Redhead, M. (1980). Models in Physics. *The British Journal for the Philosophy of Science*, 31(2), 145-163.

Rice, C. (2018). Idealized models, holistic distortions, and universality. *Synthese* 195(6), 2795-2819.

Rice, C. (2019). Models Don't Decompose That Way: A Holistic View of Idealized Models. *British Journal for the Philosophy of Science,* 70(1), 179-208.

Roca-Royes, S. (2011a). Conceivability and *de re* modal knowledge. *Noûs* 45 (1), 22-49.

Roca-Royes, S. (2011b). Modal Knowledge and Counterfactual Knowledge. *Logique et Analyse* 54, 537-552.

Roca-Royes, S. (2017). Similarity and Possibility: An epistemology of de re possibility for concrete entities. In B. Fischer & F. Leon (Eds.), *Modal Epistemology After Rationalism* (pp. 221-245). Dordrecht: Springer.

Salis, F. (2020) Learning through the Scientific Imagination. *Argumenta* 6(1): 65-80.

Sjölin-Wirling, Y. (forthcoming). Extending Similarity-based Modal Epistemology with Models. *Ergo*

Sjölin Wirling, Y. and Grüne-Yanoff, T. (2021). The Epistemology of Modal Modeling. *Philosophy Compass, 16*(10), e12775.

Stuart, M. T. (2020). The productive anarchy of scientific imagination. *Philosophy of Science*, *87*(5), 968-978.

Sugden, R. (2000). Credible worlds: the status of theoretical models in economics. *Journal of Economic Methodology* 7(1), 1-31.

Sugden, R. (2011). Explanations in search of observations. *Biology & Philosophy*, *26*(5), 717-736.

Tahko, T. (2012). Counterfactuals and modal knowledge. *Grazer Philosophische Studien* 86 (1), 93–115.

Verreault-Julien, P. (2019). How could models possibly provide how-possibly explanations? *Studies in History and Philosophy of Science Part A, 73*, 22-33.

Viswanathan, V., Atilola, O., Goodman, J., & Linsey, J. (2014, October). Prototyping: A key skill for innovation and life-long learning. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings* (pp. 1-8). IEEE.

Viswanathan, V., & Linsey, J. (2012, October). Physical modeling in design projects: Development and testing of a new design method. In *2012 Frontiers in Education Conference Proceedings* (pp. 1-6). IEEE.

Weisberg, M. (2013). *Simulation and Similiarity: Using models to understand the world*. Oxford: Oxford University Press.

Weisberg, D. S. (2019). Is imagination constrained enough for science? In Levy, A., & Godfrey-Smith, P. (Eds.) *The scientific imagination* (pp. 250-261). New York: Oxford University Press.

Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Oxford University Press.

Williamson, T. (2013). Knowing by Imagining. In Kind, A., & Kung, P. (Eds.). (2016). *Knowledge through imagination* (pp. 113-123). Oxford: Oxford University Press.

Wilson, A. (2020). *The Nature of Contingency*. Oxford: Oxford University Press.

Woodward, J. (2005). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.

Yablo, S. (1993). Is conceivability a guide to possibility? *Philosophy and Phenomenological Research,* 53 (1), 1-42.