Ergo AN OPEN ACCESS JOURNAL OF PHILOSOPHY

Extending Similarity-Based Epistemology of Modality with Models

YLWA SJÖLIN WIRLING
University of Gothenburg, University of Manchester

Empiricist modal epistemologies can be attractive, but are often limited in the range of modal knowledge they manage to secure. In this paper, I argue that one such account—similarity-based modal empiricism—can be extended to also cover justification of many scientifically interesting possibility claims. Drawing on recent work on modelling in the philosophy of science, I suggest that scientific modelling is usefully seen as the creation and investigation of relevantly similar epistemic counterparts of real target systems. On the basis of experiential knowledge of what is actually the case with the models, one can draw justified conclusions about what is *de re* possible for the target systems.

1. Introduction

Epistemologists of modality are increasingly concerned with heeding anti-exceptionalism (Williamson 2007: 136), according to which knowledge of objective modality should be construed as resulting from ordinary cognitive capacities, requiring no special resources. This recent turn has resulted in a number of *empiricist* modal epistemologies. They hold that modal knowledge ultimately rests on experiential knowledge. Existing empiricist accounts often focus on mundane examples of modal knowledge, typically concerning uncontroversial, "everyday" possibilities (e.g., "This table could possibly break"), but leave open how more interesting and cutting-edge modal claims might be justified. On the other hand, it is widely accepted that scientists often make modal claims, many of which are not of the mundane, everyday variety. Therefore, whether (and if so, what kind of) an account in keeping with the empiricist turn can be given of

Contact: Ylwa Sjölin Wirling <ylwa.sjolin.wirling@gu.se>

https://doi.org/10.3998/ergo.2246

how, for example, possibility claims of primarily scientific interest can be justified, should be an important question for modal epistemologists.

Recently, philosophers of science have emphasized that *scientific models* are used to draw modal conclusions, and while epistemologists of modality have occasionally recognised this (e.g., Nolan 2017: 23–25), there is no developed account of what the role of models is supposed to be and in virtue of what they can play this part.[1] In this paper, I offer a way of extending an attractive and plausible version of modal empiricism—Sónia Roca-Royes's similarity-based account—to cover scientifically interesting possibility claims, crucially appealing to scientific models. In particular, my suggestion is that manipulating and studying scientific models is a way of acquiring experiential knowledge of what is actually the case with certain individuals (the models), and on the basis of that knowledge we can draw conclusions about what is *de re* possible for the targets of interest. This inference is justificatory, when it is, in virtue of the fact that models are *epistemic counterparts* of their targets. Thus, my proposal represents one (although very likely not the only) way in which scientific models can be relevant to modal justification.

2. Similarity-Based Modal Empiricism

In this section I outline Roca-Royes's (2017) similarity-based, empiricist epistemology of *de re* possibilities for concrete objects. Hers is not the only candidate modal empiricist account,[2] nor is it the only similarity-based account of possibility knowledge[3], but it is intuitively plausible and arguably among the more detailed empiricist accounts on offer.[4]

Roca-Royes's story is that we can learn that it is possible for *a* to be F by consulting our knowledge of what is actually the case with distinct entities like *b* and *c*, which are relevantly similar to *a*. For example, you know that your (currently non-broken) smartphone screen can crack because you know that other smartphone screens have actually cracked; we are justified in believing that John F. Kennedy could have died of a heart attack on the basis of our knowledge that other human beings have actually died of heart attacks; I am justified in my belief that it is possible for the young horse I am looking to buy to jump a 1.3m

1. Fischer (2017) accounts for scientifically interesting possibility claims partly by appeal to models, but his sense of 'model' is different from what I have in mind. I clarify this in Section 4.

2. For some others see, e.g., Bueno and Shalkowski (2014), Leon (2017), Strohminger (2015).

3. E.g., Dohrn (2019) also assigns a large role to similarity in modal epistemology, and several others appeal to analogical reasoning more generally.

4. My aim is to extend the reach of this theory, which if successful should strengthen the case for it, but resolving other issues it faces will have to be left for another occasion.

fence with a rider, because I know many other horses of the same breed and size actually do so on a regular basis.

In more detail, the justification of a *de re* possibility claim like "It is possible for this horse, Joey, to jump a 1.3m fence" boil down to a lot of non-modal *de re* knowledge of what is actually the case with a number of actual individuals.[5] For instance, I know that Clover is a horse of a certain breed and size; that Milton is a horse of a certain breed and size; that Clover has actually jumped a 1.3m fence; that Milton has actually jumped a 1.3m fence. These are all empirically accessible facts, and my knowledge of them is based on experience. Moreover, I also have a lot of (quite trivial) modal knowledge for free—for example, it is *possible* that Milton jumps a 1.3m fence—since I am aware that actuality entails possibility. Now, with the help of enough such experiential knowledge of horses' actual properties and some inductive reasoning, I am justified in believing a principle of the form $\forall x(F(x) \rightarrow \Diamond G(x))$—in this case, for example, for all horses of a certain breed and size it is possible that they jump a 1.3m fence. Since I also know that Joey is a horse of a certain breed and size, I am justified in concluding that it is possible that he jumps a 1.3m fence even if I also know that he, in fact, has not done so.

2.1. *Epistemic Counterparts*

A lot of work is done by the fact that Joey is *relevantly similar* to Milton and Clover, whom I know have jumped 1.3m. Roca-Royes calls such relevantly similar (to the target individual) entities *epistemic counterparts* (2017: 226). Generally speaking, one is justified in concluding that *a* is possibly F, on the basis of one's knowledge that *b* is F, just in case *b* is an epistemic counterpart of *a*.

What, more exactly, does the epistemic counterpart relation involve? The notion of a counterpart is probably familiar to many from David Lewis's (1986) use of it to account for *de re* modal truths within his concrete modal realism. In that context, the counterpart relation is put to *metaphysical* use, in the sense that it is the obtaining of a counterpart relation to some distinct individual *b*, which makes it the case that so-and-so is a *de re* possibility for *a*. For instance, it is possible that Joey jumps a 1.3m fence just in case Joey has a counterpart, Joey*, in another possible world, and Joey* *in fact* jumps 1.3m fences. The epistemic counterpart relation is different from the modal counterpart relation in that it does *not* ground *de re* modal truths about the target individual. Assuming it is indeed true that Joey could possibly jump a 1.3m fence, this is not made true by

5. This is *a* way of justifying modal claims—Roca-Royes's leaves open whether there are also other ways.

the fact that Milton and Clover have actually done it. Instead, the fact that Milton has done it *informs us* about the (wholly independent) fact that Joey could possibly do it. More generally, the fact that *b* is F does *not* make true that *a* is possibly F—it informs us about the fact that *a* could possibly be F, and it can so inform us because *b* is an epistemic counterpart of *a*. It is the obtaining of the epistemic counterpart relation(s) between *a* and *b*, which accounts for *why we can infer* modal truth(s) about *a* on the basis of what we know about *b*.[6] Another difference is, as Roca-Royes takes care to stress, epistemic counterparts are *worldmates* with the target individual.

But the epistemic modal counterpart relation shares with Lewis's metaphysical modal counterpart relation the following two characteristics. First, it is a relation of *similarity*, that is, a counterpart relation obtains in virtue of a similarity between its relata. But—second—similarity itself is not enough. 'Similarity' is flexible to the point of vacuity: everything is similar to everything else in *some* respect or other. Whether *b* is a counterpart of *a*, in a given context, in virtue of a certain similarity, depends partly on our interests and intentions (cf. Divers 2002: ch. 8; Lewis 1986). So, Clover is an epistemic counterpart of Joey in virtue of the fact that they are similar in certain respects and to certain degrees *that are picked out* by the context of inquiry.

Let's consider this in a bit more detail. Entities are the way they are, objectively speaking—they have certain properties, and this is generally independent of context and interests. For instance, Clover and Joey each have certain properties, some of which they share—they are both mammals, horses, Irish hunters, and chestnuts, for instance—so these two individuals are objectively similar in those respects. In virtue of their similarities, Clover is an epistemic counterpart of Joey with respect to a range of modal questions we might be interested in, but not with respect to all such modal questions. It is the prospective possibility we are wondering about that determines which similarities are *relevant* and thus whether the epistemic counterpart relation holds. Despite their many similarities that are relevant to inquiring into what is possible for Joey, there are clearly several prospective possibilities of Joey's we might be wondering about, with respect to which Clover is *not* an epistemic counterpart of Joey. For example, Clover, who is a mare, is not an epistemic counterpart of Joey's with respect to whether it is possible for Joey, who is a stallion, to give birth to a foal. Therefore, if I conclude that it is possible that Joey gives birth to a foal, from my knowledge that Clover has actually given birth to a foal, my belief will not be justified.

6. It should also be noted that while it follows from Lewis's use of counterpart theory that the truth-value of *de re* modal claims vary with context, the notion of an epistemic counterpart has no such consequences.

In the current kind of context, where the aim is to account for how one might obtain *de re* possibility knowledge about some entity *a* starting from knowledge of what is actually the case with other individuals, the epistemic counterparts of *a* will be distinct individuals that actualize the corresponding possibility *and* are similar to *a* in respects relevant to the possibility in question.[7] That is, Clover is an epistemic counterpart of Joey in the jumping case if she has actually jumped a 1.3m fence (in contrast to a lot of other similar horses of the same breed and size who have not, and are therefore *not* Joey's epistemic counterparts with respect to this question) *and* she is similar to Joey in respects relevant to, for example, jumping capacity.

2.2. *Epistemic Counterparts and the Scope of Modal Knowledge*

According to modal empiricism, what one knows about nonactual possibilities depends on what one knows about what is the case with actual individuals. Roca-Royes intentionally does not specify exactly how that knowledge is acquired—in addition to direct and indirect observation, "ampliative methods of potentially all sorts" (2017: 230) will be involved—but at bottom the material for the ampliative reasoning is supposed to be empirical. A key virtue of the theory is that it unproblematic and uncontroversial that we possess ample knowledge of actuality.

But the apparent flipside is that the account is limited in the range of possibility knowledge it manages to secure for us. Obviously, what we are justified in believing to be possible depends on the range of our experience, and as wide as that is, it is undeniably limited in many respects. For instance, let's say I wonder about whether it is possible for Joey to jump a 2.55m fence, but I have no knowledge to the effect that any relevantly similar individuals have jumped 2.55m fences. What could I do if I wanted to find out? I could actively try to find a relevantly similar actual case. But suppose there are none.[8] Then it seems I should remain agnostic about whether or not it is possible for Joey to jump a 2.55m

7. If Joey and Tempest are identical twins and I know that it is *impossible* for Tempest to jump 2.5m, I am presumably justified in concluding that it is impossible for Joey too. In this case, Tempest is plausibly Joey's epistemic counterpart with respect to a possibility without realizing the possibility in question. However, in this case one is already in possession of some modal knowledge—of a metaphysical impossibility, no less—that is very different from the trivial possibility knowledge, inferred from actuality, involved in Roca-Royes's similarity-based account. Note that we cannot appeal to our knowledge of Tempest and the mere fact that he has *not actually* jumped a certain height, in establishing that is *not possible* for Joey to jump that height. Thanks to an anonymous referee for pressing me on this.

8. The current world record holder Husao ex-Faithful jumped 2.47m.

fence.[9] Or consider scientists who wonder about whether a particular spread of infection scenario is possible for a new virus in a certain human population. The virus has not previously been observed in human beings, nor is there data about widespread transmission of any similar virus. Presumably this epistemic situation—where there aren't any relevantly similar actual scenarios to study—is, given the current account, one where the scientists are justified in believing neither that it is possible *nor* that it is not possible.

We can express this in terms of epistemic counterparts. The more (potential) epistemic counterparts of *a* we have cognitive access to, the more we can know about the nonactual possibilities of *a*. But if *a* does *not* have any relevant epistemic counterparts with respect to a prospective possibility, I will typically not be able to form a justified belief either way.

Roca-Royes explicitly aims to account primarily for (some of) the uncontroversial, "everyday" modal knowledge that no one but the radical sceptic would deny that we have. She explicitly (2017: 223–24) pushes certain limitations as a virtue of modal epistemologies. They should first and foremost explain the kind of modal knowledge that we undeniably have. If it would just as easily grant (or require, as a basis for the everyday modal knowledge) that we know high-octane philosophical modal claims about which there is widespread disagreement, such as "The laws of nature could have been different from what they are", that should make us suspicious. [10]

Even if failing to explain extraordinary modal knowledge is virtuous, it is much less clear that it would be virtuous to fail to explain knowledge of scientifically interesting modal claims. Because beyond what we undoubtedly *know*, there are many possibilities that we are interested in finding out about and that seem within our reach: modal claims that are not completely uncontroversial, but that one can nevertheless be justified in making. Many of these concern possibilities for individual entities for which there are no relevant epistemic counterparts that actualize the corresponding state(s) of affairs that are importantly dissimilar to what we have so far experienced. Presumably, many (although not all) scientifically interesting modal claims are like this. Think again of the spread of infection scenario, for instance, or of many other cases like it: scientists have wondered (and in some cases still wonder) about whether certain climate scenarios are possible on our planet, about whether the genetic evolution of life on Earth could have been based on something other than RNA/DNA, about whether it is possible for racial segregation to occur in populations without racist preferences, about whether it is possible to resurrect the woolly mammoth using elephant

9. I won't rule out that there may be cases where an impossibility conclusion is justified on inductive grounds, but Roca-Royes similarity-based account is explicitly silent on impossibility knowledge.

10. See also Leon (2017: 259–60).

DNA, about whether it is possible for a particular naturally occurring pattern distribution to be caused by certain stochastic mechanisms, about whether it is possible that implementation of an engineering plan for securing more freshwater will result in certain disturbances of the ecosystem in the San Francisco Bay area. In these and other cases, scientists arguably cannot acquire the relevant possibility knowledge by studying actual epistemic counterparts of the targets in question. Sometimes because there clearly aren't any, sometimes because they do not or cannot (in practice or principle) know whether there are any, and sometimes because it is just not practically feasible to study the potential epistemic counterparts in enough detail. Yet scientists don't throw up their hands at these prospective possibilities, despite the fact that they are lacking experiential knowledge of the actualization of relevantly similar scenarios. Indeed, they often appear to be making justified claims about such possibilities.

Others have picked up on this limitation too. For instance, Dohrn (2019: 2468) observes that "limiting the use of analogical reasoning to actual things as Roca-Royes does leads to implausible limitations" considering what possibilities we can know about. I will presently put forward a suggestion for how to extend Roca-Royes's similarity-based modal empiricist account of *de re* possibility knowledge to cover the justification of more interesting possibility claims, which appear to have no actual epistemic counterparts.

3. The Proposal Stated

In Roca-Royes's own words, the method she outlines fails when it does because of the "impossibility of gathering empirical evidence *in favour of*" the claim in question (2017: 239).[11] I suggest that what one can do—indeed, what is often done—in the face of a lack of epistemic counterparts with respect to a certain prospective possibility, is to *create* the relevant epistemic counterparts. In a sense, the proposal picks up Dohrn's point above about the problematic limitation to "actual things".

I suggest we view the construction of *scientific models* as the creation of epistemic counterparts of certain actual systems, the possibilities of which we are interested in. Manipulating and studying models is a way of acquiring experiential knowledge of what is actually the case with individuals—the models—and on the basis of that knowledge we can draw conclusion about what is possible for the targets we are interested in. This inference is justificatory, when it is, because the models are epistemic counterparts of their targets. As above, the epistemic counterpart relation depends on our interests, but holds in virtue of

11. See also Fischer (2017: Sections 1.3 and 1.4).

an underlying relevant similarity between model and target. In the next section I outline this proposal in more detail.

4. Scientific Models as Epistemic Counterparts

Philosophers of science and others use the term 'model' in different ways. One use is associated with the semantic view of theories (Suppe 1974; Suppes 1960; van Fraassen 1989), according to which all scientific theories are sets of models. The sense of 'model' there, is the sense at work in meta-logic and set theory: models provide the semantics for the claims made by a theory. While the semantic view has many virtues as an account of how scientific theories are structured and individuated, this is not the sense of 'model' I am interested in.

Instead, I have in mind what Godfrey-Smith (2006) calls *model-based science*, which is a non-obligatory but common *strategy* that scientists employ, where they build, study, and manipulate models (for other important work highlighting modelling in the relevant sense, see, e.g., Weisberg 2013 and Giere 1988; 1999). The model-based strategy is more pervasive in some scientific fields and/or sub-fields than in others, but is undeniably a central part of contemporary scientific practice.

Models, in this sense, are a much more heterogeneous bunch than the logician's models: they range from concrete scale models, across highly complex computer simulations of the earth's climate system, to relatively simple mathematical models of population dynamics or market mechanisms. Importantly, they are often at least partially independent of theories and data—and this is often crucial to their function (see, e.g., Morgan & Morrison 1999). Models are constructed, artificial systems that are produced for the purpose of studying certain scientific problems. Sometimes models are studied merely for their own sake, but often they are used as stand-ins for real world objects and systems. One draws inferences about the so-called *target system* on basis of what one learns about the model. As it is sometimes put, models are routinely used by scientists as *epistemic surrogates* (Suárez 2004; Swoyer 1991) for some target of interest, and this practice is widely accepted as justificatory.

By way of illustration, consider again the spread of infection scenario from above. The scientists could find out by trying to actualise the scenario, for example, by planting the virus into the population itself and observing what happens. But that does not seem to be a very good idea—and neither does planting it into a relevantly similar actual population. What the scientists would presumably do instead is construct *a model* of the relevant scenario. On the basis of their findings—for example, if they, under the right constraints, manage to bring about a corresponding scenario in the model—they would then conclude that the spread of infection in question is possible in the real population of interest.[12] And, if the model is a good one, they would be justified in this conclusion.

More generally, while the best way to study something might, in theory, appear to be pushing and prodding *that thing* to see what happens, for a myriad of different reasons (e.g., ethical, economical, practical, legal) this can be difficult or even impossible. Then, one can instead construct a model to manipulate and study, and on basis of what one learns about the model, draw conclusions about the target—including modal conclusions. The function of epistemic surrogacy intuitively parallels (in a more general way) the role of epistemic counterparts, assigned to actual individuals on Roca-Royes's similarity-based modal epistemology: by studying what actually happens with some thing(s) (epistemic counterpart/surrogate), we can learn about what is possible for some distinct thing (target individual/system).

Now, in virtue of what can models serve this surrogacy function, when they can?[13] It is clear, as Morgan (1999: 366) puts it, that "[i]f we want to use [a model] to learn about the world, the model needs to map onto the real world". Philosophers of science differ on how this mapping relation should be spelled out, but it is common to do so in terms of similarity, which may be, for example, partial, structural, with respect to properties, behaviour, or causal mechanisms (Cartwright 1983; Giere 1999; 2010; Weisberg 2013; Godfrey-Smith 2006; Mäki 2009). This too parallels the epistemic counterpart relation, which obtains in virtue of an underlying objective similarity. Moreover, we saw that the relevant similarities are singled out by the context and our interests as inquirers. In the same vein, Giere notes that while bare claims of similarity are uninformative and unhelpful,

> in any particular context, what is said to be similar to what, in what ways, and to what degrees, can be specified. Of course, there is no unique specification. There are many possible specifications depending on the particular interests of those doing the modelling (1999: 46).

This has led Giere (2010) to formulate an "agent-based account of modelling" (see also Weisberg 2013) which reflects the importance of the epistemic subjects that *use* the model in specifying the relevant similarities. The result is a view of the successful epistemic surrogacy relation between model and target that is very similar to the epistemic counterpart relation, which also makes room for the

12. Of course, scientists often take stronger conclusions to be licensed too, e.g., about what *is* the case, or what *would* happen if so-and-so were the case. The epistemic surrogacy function is not reserved for, but arguably includes, modal conclusions. See Section 5.2 for related discussion.

13. I am concerned here with similarity as an answer to what explains *successful*—in the sense of justificatory—epistemic representation, rather than representation tout court. See Contessa (2007) for this distinction and its importance.

inquirer who seeks to *use* an entity as an epistemic counterpart of another entity, her interests and intentions, in addition to the underlying similarity. This gives the relation a direction, and determines what similarities are relevant. A model would be picked out as an epistemic counterpart with respect to a target and a prospective possibility of interest in the relevant context of inquiry in virtue of certain similarities plus the fact that the corresponding possibility is actualized in the model (more on this in 4.1).

Epistemic counterparthood also sits well with the fact of model pluralism. Models and target are not always related in a one-one manner. For example, the shell model, the liquid drop model, and the quark model are all used to study the atomic nucleus. Epistemic counterparthood is dependent on a context of inquiry, on the (modal) questions that we ask, so it makes perfect sense that there would be a plurality of epistemic counterparts—including models—for any given target individual. We already saw that while Clover is an epistemic counterpart of Joey with respect to several modal questions about Joey that we might be interested in, not so with respect to whether Joey could possibly give birth to a foal. However, there are plenty of other, distinct horses (and other male mammals) that *are*. Analogously, several complementary models can be constructed and utilised as epistemic counterparts of a given target, for different prospective possibilities.[14]

Not that although similarity-based accounts of modelling are congenial to my proposal, I am not committed to some of their stronger claims, because I am not suggesting that *all* models can function as epistemic counterparts in Roca-Royes's sense. "Inferentialists" such as Suárez (2004) urge that we should not characterise the model-target relation in terms of similarity, because similarity is not *necessary*—a model can map onto a target in virtue of something other than similarity. Suárez instead suggests that a model *m* maps onto a target *t* just in case *m* is such that it allows competent and informed epistemic subjects to draw specific inferences (e.g., analogical, inductive, or abductive) regarding *t*. But this is perfectly compatible with the analogy between epistemic counterparts and models as epistemic surrogates that I wish to draw here. Because the fact that a model "allows" such inferences is supposed to be an objective feature of the model: *m* has these "resources" in virtue of its "internal structure" (Suárez 2004: 773–75). A model can carry within itself the resources to be an epistemic surrogate for a wide range of different target systems, but it is the interests of the modeller, and the context of inquiry, which determines whether a model is in fact utilized as epistemic surrogate for a given target. And as Suárez also acknowledges, in many cases, the "resources" of the model,

14. Model pluralism sometimes raises issues of incompatibility, on which my suggestion has no bearing. For discussion, see, e.g., Morrison (2011) and Massimi (2018).

in virtue of which it can be successfully used as an epistemic surrogate, *does* consist in similarity of some form (which is singled out as relevant by the context of inquiry).

These general remarks about scientific models and how they function should hopefully persuade the reader that there is some initial plausibility to the idea of casting scientific models in the role of epistemic counterparts on an extension of Roca-Royes's similarity-based modal epistemology. In order to further support it, I will elucidate the proposal in more detail in the subsection below. But first I want to briefly highlight how my suggestion differs from that of Fischer (2017), as his *theory-based epistemology of modality* (TEM) also appeals to "models", albeit indirectly. According to TEM, a scientifically interesting modal claim is justified just in case its truth is implied by a scientific theory that is itself justified. In elaborating what it means for the truth of a modal claim to be implied by a theory, he makes use of the semantic view of theories. As noted above, the semantic view takes theories to be clusters of set-theoretic models. A scientific theory *T* of a phenomenon *X*, specifies a set of models, and this set represents all the possible states for *X*. Thus, *p* is possible according to *T* just in case some model in the set specified by *T* is a model of which *p* is true (2017: 19–21). The justification one has for believing that *p* is possible is parasitic on the justification one has for *T*.

Fischer's proposal is different from what I am suggesting. Similarity-judgements play no part in TEM, and the similarity-based modal epistemology does not require that the epistemic subject has a justified *theory* of, for example, predator-prey interactions in order to elicit modal knowledge about some actual predator and prey populations based on the Lotka-Volterra model. Of course, justified scientific theory, when available, is presumably involved in supporting judgements that certain models are relevantly similar to certain targets. But similarity-judgements can be made independent of confirmed theories. This also means that the current proposal can function as a complement to TEM. TEM presents a plausible enough story for how certain scientific modal claims are justified, but it is limited to areas of science where there are well-established theories available of the relevant phenomena. A key insight that has driven the shift in interest among many philosophers of science from scientific theory to models and the strategy of modal-based science is the fact that models can relate to theory in a variety of different ways. Many models are not models of theory, but are partly or wholly independent, prior or complementary to, theory (Frigg & Hartmann 2020; Reutlinger, Hangleiter, & Hartmann 2018). Scientists learn from models about the world—including about possibilities—even without having a justified theory of the relevant phenomena. TEM is silent on how that could work, while the extended similarity-based account is not.

4.1. Model-Target Similarity

In this subsection, I will develop the suggestion canvassed above with respect to two key questions: first, how the relata of the epistemic counterpart relation are to be understood on the proposed extension, and what it might mean to say that a model is relevantly similar to a target individual in the sense required to ground an epistemic counterpart relation

Philosophers who focus on model-based science have different views about the ontology of models and how models can be successful representations. While I would like to remain as neutral as possible on these issues, for the purpose of providing some more meat on the bones of the proposal to extend similarity-based modal epistemology with models, it will be necessary to have a more specific view to work with. In what follows, I will draw heavily on the seminal work of Michael Weisberg (2013). Apart from being extremely influential, Weisberg's account is congenial to my purposes because he spells the relation between model and world out in terms of similarity.

For Roca-Royes, the relata of the asymmetric epistemic counterpart relation are ordinary, concrete individuals: horses like Joey and Milton, or my desk and some other table or wooden artefact. On the extended similarity-based account, the relata of epistemic counterpart relations will be on the one hand target systems, and on the other hand scientific models. A target system is something in the real, actual world.[15] Target 'system' is to be understood broadly: in principle, Joey the horse can count as a target system, but scientific modal claims will typically concern phenomena of quite different types: a lynx population, an insect species, the Earth's climate system, a black hole, or the segregation in Philadelphia. The modal claims that the extended similarity-based modal epistemology seeks to justify will be *de re* possibility claims pertaining to such target systems, for example, what is possible as regards the dynamics of *that* lynx population, or the development of *that* climate system.[16]

The epistemic counterparts of the target systems will be scientific models, and I will assume that whatever models are, they are worldmates of target systems. According to Weisberg (2013), models are structures with certain interpretations. A structure might be concrete (e.g., an arrangement of tin-sheet and wire), mathematical (e.g., trajectories in a state space), or computational (e.g., an algorithm). The structure itself is not a model until it receives a certain interpretation by the scientist(s) who work with it: it is the double-helix arrangement

15. Some models may have fictional/imagined target systems (Bokulich 2016; Massimi 2019) but for current purposes we can ignore this option.

16. This is not to say the scientific models could not support other kinds of modal claims as well, or that more general modal or non-modal claims couldn't be made on the basis of the particular *de re* possibility knowledge described by this account.

of tin-sheet and wires *interpreted as* a DNA molecule that is Watson and Crick's model, and the trajectories described by a pair of non-linear differential equations *interpreted as* the correlated oscillation in size between a predator population and a prey population that is the Lotka-Volterra model.[17] An interpretation involves, amongst other things, a coordination between parts of the model and parts of the target system, for example, it specifies that a certain set of fluid dynamic equations in a climate model represents a certain climatic process, or which part of the San Francisco Bay hydraulic scale model represents the Golden Gate Bridge. It will also involve decisions about which features need to be taken into account and which can be disregarded when we consider how similar a model is to its target. Weisberg distinguishes between two kinds of properties or features that a model has, and that may be included in this set: *attributes*, which are *states* of the model, and *mechanisms*, which are the *transition rules* that bring about states in the model (2013: 145). What goes in this set of relevant features depends on the context of inquiry, and on the goals that the scientists' wish to use the model for (e.g., prediction and explanation may call for different features to be included). When using a model to investigate possibilities, which attributes and mechanisms of the model go in this set will be what the scientists take to be relevant to the prospective possibility in question. For instance, in order for the San Francisco Bay hydraulic scale model to be an epistemic counterpart of the Bay area with respect to the possibility of a certain disturbance in the ecosystem, initial salinity levels and tidal and river flows will be among the properties in the feature set, while surface tension and elasticity will not (cf. Weisberg 2013: ch. 8). But of course, it is sometimes difficult to tell with any certainty what features are relevant to investigating a given possibility, especially when data and/or theory are scarce.

Target-directed modelling has three steps: construction of the model, analysis of the model, and fitting the model to the target system. Choice of structure, and interpretation of the structure, are part of the construction. Analysis of the model involves learning about the model, for example, how it works, how it responds to manipulation, and how its parts depend on each other. Fitting the model to the target involves assessing it for relevant similarity (or dissimilarity) and, when possible, amending or calibrating the model to achieve a higher degree of relevant similarity for the scientific purpose at hand. As Weisberg notes (2013: 74–75), while conceptually separate, the steps often overlap and scientists may move back and forth between them.

17. This means that one and the same structure can be part of several models, as when the state space trajectories described by Volterra's equations are interpreted in terms of correlated oscillation of the workers' wage share of the national income and the employment rate.

It is worth pausing to see how this compares to the process described by Roca-Royes's similarity-based modal epistemology. There, the idea is that epistemic subjects draw on their existing empirical knowledge of what is actually the case with relevantly similar individuals (e.g., Milton and Clover), in order to find out about whether something or other is possible for a given target individual they are interested in (Joey). The rationale for extending the similarity-based account with models was precisely *a lack* of existing empirical knowledge about relevantly similar actual cases—so in an instance of active inquiry, aiming to extend one's knowledge, one turns to models. One can construct a model from scratch with the explicit purpose of finding out about a particular prospective possibility, but in reality, one will more likely adopt, perhaps with adjustments in interpretation, some existing model(s) of the relevant type of phenomenon already in use. In order to establish the model as an epistemic counterpart of the target system, one needs to establish that the model is relevantly similar to the target. But one also needs to make the model actualize the relevant possibility that one is interested in. Analysis of the model can here be seen as roughly analogous to my actively attempting to acquire more knowledge about potential epistemic counterparts (e.g., Milton and Clover) of the target (Joey), including knowledge about whether they actualize the relevant possibility. Fitting the model to the target can be seen as analogous to settling to what extent Milton and Clover are relevantly similar to Joey. But—and here is an important advantage with models—it might also involve adjusting the model so that it becomes more similar in relevant respects, so that it can better serve as an epistemic counterpart.

a is an epistemic counterpart of *b* with respect to a prospective possibility, say whether *b* could be *F*, just in case *a* is relevantly similar to *b* and *a* is F. I claim that models can be epistemic counterparts, in something like this sense, of target systems. Following Weisberg and others, I assume that models can be evaluated for similarity with their targets, by examining to what extent they have or lack the attributes or mechanisms specified in the feature set. But models are evidently very different things from target systems. This is most obvious in the case of mathematical and computational models—a computer simulation of the Earth's climate system is not a climate system—but it holds for concrete models too: Watson and Crick's tin-sheet-and-wire double helix model of the DNA is not itself a DNA molecule, and even a scale model of an aeroplane wing is not an aeroplane wing. Models and targets, being different kinds of things, instantiate different kinds of properties. Therefore, the sense in which they can be similar cannot be understood in terms of literal property sharing, as we might be inclined to understand the similarity between Joey and Milton.

This raises two related questions: how can a model and a target be relevantly similar if they do not instantiate the same properties, and what does it mean to say that a model actualizes a property that we wonder whether the target system could have? While there is no way in which the issue can be given full treatment in this paper, I will offer two considerations that I hope will be helpful.

First, as regards the question of how we could even begin to compare a model and a concrete target, Weisberg writes:

> [H]ow can mathematical and computational models be compared to concrete targets? What possible similarities do they have? I think this objection can be easily resolved. Mathematical and computational models, as well as concrete models in some cases, are compared to mathematical representations of targets, not the targets themselves. [. . .] In simple, dynamical models, the mapping is such that the major determinable properties (e.g. species abundance, pressure, time, temperature, etc.) of the target are mapped to dimensions of a state space. Now one interpreted mathematical object can be compared to another, and we avoid problems about comparing mathematical properties to concrete properties (2013: 95).

Obviously, the mathematical representation of the target will need to be constructed with a close eye to what one knows about the target. But this is not different in kind from judgements of relevant similarity more generally: in order to judge competently whether Milton is relevantly similar to Joey, I need to know things about the target, Joey. Of course we might need more advanced knowledge, but that is as it should be—science is more demanding than everyday inquiry. Exactly how the relevant background knowledge is acquired, and how much of it there is to draw on, will differ greatly between scientific disciplines and cases.

Second, it will be useful to follow the advice of Khosrowi (2020) who has recently proposed an improvement of Weisberg's account of model-target similarity, in taking a pluralistic approach to what it might mean for models and targets to be similar. Clearly, it is a huge and complex question how actual features are best scientifically represented. For one, the type (e.g., concrete, mathematical, computational) of the model will call for different modes of representation, and a given attribute or mechanism will be represented differently in a computer simulation and in a scale model. For another, different kinds of properties (of targets)—structural, causal, categorical—plausibly need to be represented in different ways. Therefore, as Khosrowi points out, different ways of comparing model and target will be suitable depending on the kind of property: for example, structural similarity might be usefully understood in terms of (partial) isomorphism, while qualitative and categorical features such as causal relationships, morphological features, shapes, and so on are better understood in terms of a similarity between the features themselves[18] (rather than literal sharing of properties). Thus, models may be similar to their targets in virtue of quantitative closeness, isomorphism, similarities of features, and possibly other relations—and in some cases of extremely high-fidelity models they might literally share properties. These are not mutual exclusives, a model-target pair can be relevantly similar in several different ways.[19]

Having said this in response to the question of how a model and a target can be relevantly similar despite being different kinds of things, we can now respond to the latter question of what it means for a model to actualise a possibility in the following way: We are interested in whether target *a* could be F, and will thus be looking for a model that is relevantly similar to *a* and is also F. To say that "the model is F" is just to say that in addition to having certain attributes and/or mechanisms that are relevantly similar (or, e.g., structurally isomorphic, or quantitively sufficiently close to) to the properties and/or causal structure of *a*, the model also has an attribute or mechanism that is relevantly similar to (or structurally isomorphic, or quantitively sufficiently close . . .) to F.

It bears pointing out that while obviously very important, this issue is a general one for accounts of successful epistemic model representation in the sciences. It faces *any* account according to which a similarity relation between model and target sometimes underwrites the epistemic legitimacy of conclusions about the world (whether modal or non-modal) based on scientific modelling. Notably, the issue doesn't go away even if Suárez and other critics of pure similarity–account of models are right that the epistemic success of several models is underwritten by something other than model-target similarity. As long as similarity in some form or other underwrites the success of *some* models—which is a widely shared assumption—the question remains. There is no reason to think that it arises in any particularly intractable form for my suggestion that models can function as epistemic counterparts that inform us about the *de re* possibilities for targets.

5. Two Challenges

In this final section, I address two challenges to the claim that I put forward and motivated above. The first is a worry about whether the extended similarity-based

18. For an early version of the notion of similarity of features/properties (rather than similarity as property-sharing), see Hesse's (1963) work on the role of analogy in modelling and her discussion of different types of similarity relations that can support analogical reasoning. She suggests, amongst other things, that entities can be similar in virtue of having properties that in turn resemble each other. One of her own examples is that that sound and light are similar in this sense, because echoes are similar to reflection, loudness to brightness, pitch to colour, and so on.

19. For work on how to deal with model-target comparisons given fictionalism about models see Salis (2016).

modal epistemology is still a form of modal empiricism, or whether it collapses into something on a par with more traditional modal epistemologies based on, for example, conceivability. The second worry concerns how my proposal here squares with some existing work in the philosophy of science on modelling that aim to establish modal knowledge.

5.1. The Challenge from Empiricism

I have presented Roca-Royes's similarity-based modal epistemology as an empiricist modal epistemology. As I noted in 2.2, we may be reasonably liberal about just how the non-modal knowledge of what is actually the case with the epistemic counterparts is acquired, but at bottom it is supposed to be empirical. That is allegedly what sets accounts like the similarity-based one apart, in a positive way, from more traditional modal epistemologies where justification for possibility claims comes from, for example, conceivability or intuition, so this objection is important to address. But it might be complained that knowledge of what is going on with models is not *experiential* knowledge in any sense of the word that would appease the empiricist-minded modal epistemologist. Indeed, don't scientists use models precisely when they lack access to certain empirical knowledge?

To be clear, my claim is not that models afford empirical knowledge of anything but *themselves*. Nevertheless, I do wish to claim that the knowledge scientists gain about the model should often be considered empirical enough for a modal empiricist, in the sense that it evades worries that modal empiricists have about modal epistemologies based on imagination, conceivability, or intuition. In showing that the extended similarity-based account is better off than modal epistemologies with a more rationalist flavour, I will call attention to two aspects shared by many instances of scientific modelling: the *role of sensory experience* in modelling and the ways in which structures (whether concrete or mathematical/computational) *constrain* modelling. As a caveat, let me note that I will not be concerned with converting those who think imagination/conceivability is a perfectly good route to possibility knowledge. My aim is solely with diffusing worries that a modal empiricist might have.

When working with a concrete model like Watson and Crick's DNA model or a hydraulic scale model of the San Francisco Bay, there is a clear sense in which scientists that manipulate, analyse, and tweak the model gain perfectly ordinary experiential knowledge—based on observation, measurement, etc.—of what is actually the case with the model. But consider the purely mathematical models that abound in biology, economics, and beyond, like the Lotka-Volterra model of predator-prey population dynamics. How do scientists learn about such a model? If the answer is just that they are explored in the imagination, or "in the mind's eye", this sounds suspiciously like the methods modal empiricism seeks to move away from, and not much like experiential knowledge at all. Indeed, there are accounts of modelling in the literature on which this worry is really accentuated. For instance, "direct" fictionalists like Toon (2012) and Levy (2015) both hold that modelling just is imagining about real-world targets. On this view, there is no epistemic surrogate postulated—no model system—"between" the scientists' imagination and the world. Scientific modelling then seems to be nothing more than a kind of thought experimenting.

This view of scientific modelling departs from the epistemic surrogate approach that is more common, and which I have assumed to be correct throughout this paper. Indeed, direct fictionalism has been subject to extensive criticism from non-fictionalists (e.g., Weisberg 2013: ch. 4.4) as well as from other ("indirect") fictionalists (e.g., Frigg & Nguyen 2016; Salis in press), partly precisely because they fail to explain the aspects of scientific modelling that sets it apart from mere imagining.

Gaining knowledge of, and from, models is typically different from merely exploring one's imaginings in a number of ways, and as, for example, Knuuttila (2017) has argued, not doing justice to this fact is a serious drawback of fictionalist accounts that put too much emphasis on the scientist's imaginings.

We can use Weisberg's useful highlighting of three ways in which scientific models differ from thought experiments to draw attention to the ways in which modelling is different from—and thereby often an epistemically superior tool to—imagining. First, modelling unlike thought experimenting requires one to make one's assumption explicit.

> Our imaginations are very flexible and we can construct an imagined scenario from a very minimal script. But when one has to derive an equation, build something out of plastic, or write a computer program, this kind of vagueness is not allowable. Programs won't compile, equations won't be derived, and plastic models will fall apart if details are left unspecified. (Weisberg 2016: 280)

Consider the following objection to imagination-based modal epistemologies. It might seem easy to imagine a scenario in which there is a naturally purple cow, indeed, one can do so without knowing anything about genetics. But as, for example, Van Inwagen (1998) has famously complained, imaginings that lack relevant details spelling out just how the relevant scenario comes about hardly provides justification for believing that the scenario in question—for example, that cows could be naturally purple—is possible. The fact that model assumptions must be properly spelled out and not left unspecified indicates that modelling is less vulnerable to this kind of objection.

Second, and relatedly, unlike thought experimenting modelling significantly reduces variation between epistemic subjects. If you and I are to imagine *a* being *F* in order to see whether that is possible, we may well fill this out in very different ways and therefore come to different conclusions, or we may imagine the same scenario but disagree whether it is a case in which *a* is *F*. Models, in contrast, will deliver a determinate result and everyone can examine the details that generated it. Modelling will not wipe out disagreement—there may be disagreement about what properties should go in the relevant feature set, or how a given aspect of a target is best represented in a model. But because models, including their driving assumptions, are intersubjectively available, can be compared and examined and worked on by different subjects, these issues seem more tractable than when we disagree about what would happen in a hypothetical scenario that we each merely imagine. Indeed, the agreement and efficient coordination of scientists using the same model is considered an important fact for modelling accounts to explain (Knuuttila 2017; Weisberg 2013: §4.4).

Third, while imagination is capable of roaming far and wide in many respects, it is also limited in many others. We may allow that mere imagination can be a good way of exploring nonactual scenarios that are comparatively simple and close to our experiences, and that these cases won't call for being mathematically, concretely or computationally modelled. But scientifically interesting modal claims will typically be vastly more complex, and involve high-dimensional spaces, non-linear systems, macro-patterns emerging as a result of massive aggregation over agents, probabilistic relationships, complicated dependencies, etc., and modelling can in many cases help us transcend the cognitive limitations that prevent us from assessing them in the imagination.

Relatedly, a relevant difference between models and imaginings is that even when models are partly constituted by abstract structures, they are not just studied in the imagination but via external representational means, and several authors have presented detailed case studies indicating that the (nature of the) external representational means are important to the epistemic function of scientific models. As pointed out by Vorms (2011: 290) and Knuuttila (2011; 2017), and as Weisberg also recognises (2013: 31–39), for models that have abstract structures, *model descriptions* are crucial for our epistemic access to those models *at all*.[20] Model descriptions are the various ways in which models are presented to their

20. "Indirect" fictionalists are careful to point out that the games of make-believe (cf. Walton 1990) that scientists engage in when they model are partly constrained by the "props" that prompt the imaginings, where the "props" in question are model descriptions. They also require that props can in principle be "[a]nything capable of affecting our senses" (Salis & Frigg 2020: 35), thereby highlighting the importance of the material dimension for epistemic access.

users, as, for example, pictures, drawings, graphical displays, diagrams, words, or equations.[21] Several authors have presented detailed case studies indicating that the external representational means are important to the epistemic function of scientific models: for example, Vorms (2011) shows that different representations of the simple pendulum model enable users to perform different inferences about its target, Morgan (2012: ch. 3) has a similar discussion of how different visual representations of mathematical models matter to what their epistemic contribution is, and Humphreys (2004: ch. 4) describes at length the importance of computer simulations in human epistemic access to certain results and upshots of theoretical, purely mathematical models. The independent structure and the material dimension of models—whether directly by way of its structure, or by way of the model description—is thus both enabling and constraining. On the one hand, models can help epistemic subjects assess prospective possibilities that the imagination cannot afford reliable access to. On the other hand, in contrast to how an imaginer is free to develop a scenario in whatever way she likes, what can and cannot happen in the model system is not "up to" the modeller since models in virtue of their structure (both formal and material) are more like autonomous objects (cf. Gelfert 2016: ch. 5; Knuuttila 2011; 2017; Knuuttila & Koskinen 2021). Note—in relation to the point about making assumptions explicit and how that might help with an objection like van Inwagen's—it is not that we could not visually imagine (rather than, e.g., see on a computer screen) the result of, for example, a complex computational model once we are aware of it. The point is that access to this model result, and justification for accepting it as genuine and relevant to, for example, a prospective possibility, comes from working with the model.

I will not insist that the above is true of all models—some scientific models may indeed be equivalent to thought experiments. Perhaps *de re* possibility-inferences based on them are sometimes justified, perhaps not—there is no need to take a stand on that here. But in those cases, it would be wrong to say that modellers gain empirical knowledge of what is actually the case with the model when they study it, and those models would not be suitable to play the role of epistemic counterparts in the sense required by similarity-based modal empiricism.

5.2. *The Challenge from Modal Modelling*

As we saw in Section 4, it is widely discussed in virtue of what a model—quite generally—fits with its target, in the sense that we can draw justified conclusions

21. Whether model descriptions are to be counted as proper parts of the models or not is an issue of some controversy. Knuuttila (2017) argues that they should, whereas Weisberg (2013: ch. 3) thinks they should not.

about the latter on the basis of the former. Relevant similarity is one paradigmatic answer to that question. I drew on that in supporting my claim that scientific models lend themselves very naturally to the role of epistemic counterparts of actual targets, from which we could draw justified conclusions about what is *de re* possible for the latter, on the basis of knowing what actually happens to the former. But here is a two-fold wrinkle with respect to this claim.

On the one hand, while scientists use models as epistemic counterparts, in some intuitive sense of that word, of actual targets, it might be objected that models are rather used in drawing stronger conclusions, for example, what *actually* is (if the concern is explanation), or will be (if the concern is prediction), the case with the target system. This is certainly true, but there is no question that models are also used for the purpose of acquiring modal knowledge, including about *de re* possibilities known to be (currently) non-actual. In recent years, a large number of philosophers of science have highlighted modelling practices that aim at drawing modal conclusions, often accompanied by detailed case studies (e.g., Batterman & Rice 2014; Gelfert 2019; Grüne-Yanoff 2013; Massimi 2019; Reutlinger et al. 2018; Rice 2019; Verreault-Julien 2019; Weisberg 2013: ch. 7). These modal conclusions may then in turn sometimes inform or figure in "stronger" claims (e.g., counterfactual information might be relevant to scientific explanation, or knowledge that *p* is possible may together with other considerations licence the conclusion that under certain circumstances *p* will be the case), but the information retrieved from the model is modal.

Closer attention to this rejoinder, however, reveals the second fold of the wrinkle. Relevant similarity, as specified by the context of inquiry, is one paradigmatic way to spell out model-target fit. But the philosophers who stress a modal interpretation of some models often do so as a way of explaining their epistemic relevance, often because a *lack* of model-target similarity appears to make them unsuited to support conclusions about what actually is the case with some target system. So the "modal move" typically rests on the assumption that retreating to possibility claims relaxes the conditions on the model-target relationship: for example, lack of similarity prevents one from drawing conclusions about actuality but we might still learn about *possibilities*. In light of this, it appears odd to claim that scientific models can be used to deliver modal knowledge in virtue of their being similar to their targets.

I am not too worried about this objection. Modal modelling will need to obey *some* constraints, in order to be justificatory, assuming there are objective facts of the matter about what is possible for a given target system. Proponents of modal modelling have not offered much of an answer yet as regards what these constraints might be.[22] Similarity-based modal empiricism, with models cast in

22. For an overview of the existing preliminary proposals see Sjölin Wirling and Grüne-Yanoff (2021) and for some critical discussion see Sjölin Wirling (2021).

the role of epistemic counterparts, suggests that relevant similarity is the answer *also* to the question of what makes a good fit for the purpose of drawing justified possibility conclusions. Or rather, it is *one* answer—perfectly compatible with there being a catalogue of other answers too. Pluralism does seem to be the right approach here, given the great diversity of scientific modelling practices, both generally and in relation to modal matters. In particular, the extended similarity-based account is an answer that is suitable when the target is a *de re* possibility claim with respect to a given actual target system—that is all I have claimed in this paper. Scientists who turn to modelling may also be interested in establishing other types of modal claims in which case a different story might be needed to spell out the justificatory route.

We shouldn't be too quick to circumscribe the relevance of the similarity-based account either. Although modal modelling interpretations are often offered in alleged contrast to, and claim to eschew, similarity-based accounts of successful epistemic representation, some of them may turn out to crucially rely on similarity-judgements anyway. For instance, Collin Rice (2018; 2019; see also Batterman & Rice 2014) argues that highly idealised "minimal" models, like the Lattice Gas Automaton model, afford modal knowledge of real-world targets (in this case, e.g., real fluid flows) despite being highly dissimilar to them in the sense of having very different properties and component parts. But in order to afford this modal knowledge, the minimal models and the target systems must be in the same "universality class", which Rice cashes out in terms of "similar patterns of macroscale behaviour" (2018: 2796). In the end then, it is *behavioural similarity* between tokens that underwrites modal reasoning from model to target.[23]

6. Conclusion

I have argued that similarity-based modal empiricism can be extended to cover interesting scientific possibility claims by casting scientific models in the role of epistemic counterparts. Learning what is actually the case with models allows scientists to draw justified conclusions about what is *de re* possible for targets, in virtue of the fact that models are relevantly similar to targets. Apart from benefiting modal empiricism, my proposal also constitutes the careful beginnings of an answer (although likely one of many possible) to the so far underexplored question of in virtue of what (some) scientific models can be good guides to (some) modal truths. There is much work left to be done, of course, not least with respect to what, if anything, can in general be said about what similarities are *relevant* to

23. For further discussion of the universality account of modal modelling see Lange (2015) and Sjölin Wirling and Grüne-Yanoff (2021).

justifying a possibility claim, how they may differ from the degree or type of similarity needed to justify an actuality claim on the basis of a model, and just how we come by knowledge about or a capacity to reliably judge about those similarities. Sorting these issues out would be an important part of further developing the proposal I am making in this paper, but the fact that there is no obvious answer to them does not threaten to undermine the suggestion as things currently stand. Moreover, these questions remain open also for the original similarity-based modal epistemology, so it is not a task particular to the extension I am proposing here.

Acknowledgements

Thanks to Till Grüne-Yanoff, Sònia Roca-Royes, Frans Svensson, and two anonymous referees of this journal for their very helpful comments. This research was funded by the Swedish Research Council, grants no. 2018–01353 and no. 2019–00635.

References

Batterman, Robert W. and Collin Rice (2014). Minimal Model Explanations. *Philosophy of Science, 81*(3), 349–76.

Bokulich, Alisa (2016). Fiction as a Vehicle for Truth: Moving Beyond the Ontic Conception. *The Monist*, 99(3), 260–79

Bueno, Otavio and Scott A. Shalkowski (2014). Modalism and Theoretical Virtues: Toward an Epistemology of Modality. *Philosophical Studies, 172*(3), 671–89.

Cartwright, Nancy (1983). *How the Laws of Physics Lie*. Oxford University Press.

Contessa, Gabriele (2007). Scientific Representation, Interpretation, and Surrogative Reasoning. *Philosophy of Science, 74*(1), 48–68.

Divers, John (2002). *Possible Worlds*. Routledge.

Dohrn, Daniel (2019). Modal Epistemology Made Concrete. *Philosophical Studies, 176*(9), 2455–75.

Fischer, Bob (2017). *Modal Justification via Theories*. Springer International.

Frigg, Roman (2010). Models and Fiction. *Synthese, 172*(2), 251–68.

Frigg, Roman and Stephen Hartmann (2020). Models in Science. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 ed.). Retrieved from https://plato.stanford.edu/archives/spr2020/entries/models-science/

Frigg, Roman and James Nguyen (2016). The Fiction View of Models Reloaded. *The Monist, 99*(3), 225–42.

Gelfert, Axel (2016). *How to Do Science with Models: A Philosophical Primer*. Springer International.

Gelfert, Axel (2019). Probing Possibilities: Toy Models, Minimal Models, and Exploratory Models. In M. Fontaine, C. Barés-Gómez, F. Salguero-Lamillar, L. Magnani, and Á. Nepomuceno-Fernández (Eds.), *Model-Based Reasoning in Science and Technology* (3–19). Springer International.

Giere, Roland. N. (1988). *Explaining Science: A Cognitive Approach*. University of Chicago Press.

Giere, Roland N. (1999). Using Models to Represent Reality. In Lorenzo Magnani, Nancy J. Nersessian, and Paul Thagard (Eds.), *Model-Based Reasoning in Scientific Discovery* (41–57). Kluwer Academic/Plenum Publishers.

Giere, Roland N. (2010). An Agent-Based Conception of Models and Scientific Representation. *Synthese, 172*, 269–81.

Godfrey-Smith, Peter (2006). The Strategy of Model-Based Science. *Biology and Philosophy, 21*(5), 725–40.

Grüne-Yanoff, Till (2009). Learning from Minimal Economic Models. *Erkenntnis, 70*(1), 81–99.

Grüne-Yanoff, Till (2013). Appraising Models Nonrepresentationally. *Philosophy of Science, 80*(5), 850–61.

Hesse, Mary B. (1963). *Models and Analogies in Science*. Notre Dame University Press.

Humphreys, Paul (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford University Press.

Khosrowi, Donal (2020). Getting Serious about Shared Features. *The British Journal for the Philosophy of Science, 71*(2), 523–46.

Knuuttila, Tarja (2011). Modelling and Representing: An Artefactual Approach to Model-Based Representation. *Studies in History and Philosophy of Science Part A, 42*(2), 262–71.

Knuuttila, Tarja (2017). Imagination Extended and Embedded: Artifactual versus Fictional Accounts of Models. *Synthese*. https://doi.org/10.1007/s11229-017-1545-2

Knuuttila, Tarja and Rami Koskinen (2021). Synthetic Fictions: Turning Imagined Biological Systems into Concrete ones. *Synthese, 198*, 8233–8250. https://doi.org/10.1007/s11229-020-02567-6

Lange, Marc (2015). On 'Minimal Model Explanations': A Reply to Batterman and Rice. *Philosophy of Science*, *82*(2), 292–305.

Leon, Felipe (2017). From Modal Skepticism to Modal Empiricism. In Bob Fischer and Felipe Leon (Eds.), *Modal Epistemology after Rationalism* (247–61). Springer International.

Levy, Arnon (2015). Modeling Without Models. *Philosophical Studies*, 172(3), 781–98.

Lewis, David (1986). *On the Plurality of Worlds*. Blackwell.

Massimi, Michela (2018). Perspectival Modeling. *Philosophy of Science, 85*(3), 335–59.

Massimi, Michela (2019). Two Kinds of Exploratory Models. *Philosophy of Science, 86*(5), 869–81.

Morgan, Mary S. (1999). Learning from Models. In Mary S. Morgan and Margaret Morrison (Eds.), *Models as Mediators: Perspectives on Natural and Social Science* (347–88). Cambridge University Press.

Morgan, Mary S. (2012). *The World in the Model*. Cambridge University Press.

Morgan, Mary S. and Margaret Morrison (1999). *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge University Press.

Morrison, Margaret (2011). One Phenomenon, Many Models: Inconsistency and Complimentarity. *Studies in History and Philosophy of Science, 42*(2), 342–51.

Mäki, Uskali (2009). MISSing the World. Models as Isolations and Credible Surrogate Systems. *Erkenntnis, 70*(1), 29–43.

Nolan, Daniel (2017). Naturalised Modal Epistemology. In Bob Fischer and Felipe Leon (Eds.), *Modal Epistemology after Rationalism* (7–28). Springer International.

Reutlinger, Alexander, Dominik Hangleiter, and Stephen Hartmann (2018). Understanding (with) Toy Models. *British Journal for the Philosophy of Science, 69*(4), 1069–99.
Rice, Collin (2018). Idealized Models, Holistic Distortions, and Universality. *Synthese, 195*(6), 2795–819.
Rice, Collin (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, Sònia (2017). Similarity and Possibility: An Epistemology of De Re Possibility for Concrete Entities. In Bob Fischer and Felipe Leon (Eds.), *Modal Epistemology after Rationalism* (221–45). Springer International.
Salis, Fiora (2016). The Nature of Model-World Comparisons. *The Monist, 99*(3), 243–59.
Salis, Fiora (in press). The New Fiction View of Models. *British Journal for the Philosophy of Science*.
Salis, Fiora and Roman Frigg (2020). Capturing the Scientific Imagination. In Peter Godfrey-Smith and Arnon Levy (Eds.), *The Scientific Imagination* (17–50). Oxford University Press.
Sjölin Wirling, Ylwa (2021). Is Credibility a Guide to Possibility? A Challenge for Toy Models. *Analysis, 81*(3), 470–478. https://doi.org/10.1093/analys/anab013
Sjölin Wirling, Ylwa and Till Grüne-Yanoff (2021). The Epistemology of Modal Modelling. *Philosophy Compass, 16*(10), e12775. https://doi.org/10.1111/phc3.12775
Strohminger, Margot (2015). Perceptual Knowledge of Nonactual Possibilities. *Philosophical Perspectives, 29*(1), 363–75.
Suárez, Mauricio (2004). An Inferential Conception of Scientific Representation. *Philosophy of Science, 71*(5), 767–79.
Suppe, Frederick (1974). *The Structure of Scientific Theories*. University of Illinois Press.
Suppes, Patrick (1960). A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences. *Synthese, 12*(2–3), 287–301.
Swoyer, Chris (1991). Structural Representation and Surrogative Reasoning. *Synthese, 87*(3), 449–508.
Toon, Adam (2012). *Models as Make-Believe: Imagination, Fiction, and Scientific Representation*. Palgrave-Macmillan.
van Fraassen, Bas C. (1989). *Laws and Symmetry*. Oxford University Press.
Van Inwagen, Peter (1998). Modal Epistemology. *Philosophical Studies, 92*(1), 67–84.
Verreault-Julien, Philippe (2019). How Could Models Possibly Provide How-Possibly Explanations? *Studies in History and Philosophy of Science Part A, 73*, 22–33.
Vorms, Marion (2011). Representing with Imaginary Models: Formats Matter. *Studies in History and Philosophy of Science, 42*(2), 287–95.
Walton, Kendall L. (1990). *Mimesis as Make-Believe*. Harvard University Press.
Weisberg, Michael (2013). *Simulation and Similiarity: Using Models to Understand the World*. Oxford University Press.
Weisberg, Michael (2016) Modeling. In Herman Cappelen, Tamar Szabo Gendler and John Hawthorne (Eds.) *The Oxford Handbook of Philosophical Methodology* (262–86). Oxford University Press.
Williamson, Timothy (2007). *The philosophy of philosophy*. Blackwell Publishing.