Received: 27 January 2021 | Revised: 4 August 2021 | Accepted: 8 September 2021
DOI: 10.1111/phc3.12775

ARTICLE

WILEY

The epistemology of modal modeling

Ylwa Sjölin Wirling[1,2] | Till Grüne-Yanoff[3]

[1]Department of Philosophy, Linguistics, and Theory of Science, University of Gothenburg, Goteborg, Sweden

[2]Department of Philosophy, School of Social Sciences, University of Manchester, Manchester, UK

[3]Department of Philosophy and History of Technology, KTH Royal Institute of Technology, Stockholm, Sweden

Correspondence
Ylwa Sjölin Wirling, Department of Philosophy, Linguistics, and Theory of Science, University of Gothenburg, Goteborg 405 30, Sweden.
Email: ylwa.sjolin.wirling@gu.se

Funding information
Vetenskapsrådet, Grant/Award Number: 2018-01353, 2019-00635

Abstract

Philosophers of science have recently taken care to highlight different modeling practices where scientific models primarily contribute *modal* information, in the form of for example possibility claims, how-possibly explanations, or counterfactual conditionals. While examples abound, comparatively little attention is being paid to the question of *under what conditions*, and *in virtue of* what, models can perform this epistemic function. In this paper, we firstly delineate *modal modeling* from other modeling practices, and secondly review attempts to spell out and explain the epistemic success conditions of modal modeling. The aim is to more clearly expose the respective justificatory strategies of these accounts, and secondly, to identify lacunae where further work is needed.

1 | INTRODUCTION

Recently, as part of philosophers' increased attention to scientific modeling, a small but active literature focusing specifically on modeling with *modal* aspects has emerged. On the one hand, scientists sometimes explicitly describe the aim or result of certain modeling practices with modal terms like possibility, dispositions, or counterfactual conditionals. On the other hand, philosophers of science have offered modal interpretations of existing modeling practices that are not overtly modal, but which more traditional philosophical accounts of modeling fail to make sense of. In both cases, the idea is that certain instances of scientific modeling deliver modal information. We will refer to these practices as *modal modeling*.

Modal modeling is a diverse phenomenon, both as regards the scientific fields, types of models and modal claims involved, and the different roles that modal modeling practices play in science more generally. The existing literature focuses primarily on documentation and analysis of specific modeling practices, although some overviews have been attempted (e.g. Gelfert, 2019). But it has not been much discussed *under what conditions*, and *in virtue of what*, scientific models can perform the epistemic function—helping to uncover modal information—that is ascribed to them. We call this the *epistemic question for modal modeling*. Our aim in this paper is to provide a more comprehensive grasp on the epistemic issue raised by modal modeling, by exposing more clearly the justificatory strategies presented in accounts of modal modeling and highlight some lacunae where further work is needed. In section 2, we delineate the phenomenon of modal modeling and briefly describe some of its forms as identified by philosophers of science. Section 3 specifies the epistemic issue that modal modeling raises, and in section 4 we review three recent attempts to tackle it. In section 5, we identify the justificatory means that these strategies rely on, and in 6 we raise some critical questions that need to be addressed in future research of the epistemology of modal modeling.

2 | MODAL MODELING

By "modal modeling" we mean modeling practices that aim at delivering *modal* information. This modal information comes in different forms: about what is possible, what would be the case under counterfactual circumstances, or potentialities of specific entities or systems. As noted already, this aim may be explicitly articulated by scientists using the model, or it may be that the modeling practice in question is, according to philosophers of science, best rationally reconstructed as delivering modal information.

This is in contrast to modeling that delivers information about what *actually* is, was, or will be the case.[1] Now, this is not to say that scientists engaging in modal modeling are not interested in the actual world. Quite to the contrary, they will often be after modal information as a means of further inquiry into, explaining, or increasing understanding of (some aspect of) the actual world. Modal modeling is characterized by an intention to deliver modal information as its immediate result, whether or not those results are then used as stepping stones to some other epistemic aim.

Philosophers often distinguish different concepts of the possible that relate differently to the actual. In particular, *epistemic possibility* concerns propositions that might be true in the actual world, in the sense that they do not contradict what is known to be actually true. Nothing known to be actually false is epistemically possible. In contrast, *objective possibility* concerns how the world could be, even if it is known not to be that way. Both objective and epistemic possibilities may be of interest to scientists, and we intend modal modeling to encompass practices aiming to probe both or either type of possibility (Sjölin Wirling & Grüne-Yanoff, forthcoming).

Modal modeling is employed in several different scientific contexts for several different purposes. The clearest examples of modal modeling are modeling practices connected with *how-possibly explanations* (HPEs). Philosophers disagree on how HPE practices are to be characterized, but most parties appear to agree that (i) models play a crucial role in supporting HPEs; and (ii) offering an HPE involves making some form of modal claim (see e.g. Bokulich, 2014; Grüne-Yanoff, 2009; 2013; Reutlinger et al., 2018; Verreault-Julien, 2017; 2019; Weisberg, 2013, chapter 7). That is, a possibility claim is made on the basis of a model result. For instance, Rohwer and Rice (2013), point out that the Hawk-Dove model (Smith and Price, 1973) supports an HPE of the phenomenon of restraint in combat between members of the same species. The model is used "to test whether it is possible even in theory for individual selection to account for 'limited war' behavior" (1973, 15). Other highly idealized and simple models in biology and economics have been argued to function this way too (Reutlinger et al., 2018). HPEs, as well as other forms of modal modeling, can contribute to a deeper scientific understanding of the studied phenomena, for instance by allowing scientists to draw correct counterfactual inferences, thereby providing an ingredient of successful scientific explanation. This is how Ylikoski and Aydinonat (2014) describe the epistemic contribution of Schelling (1971) checkerboard model. Similarly, Rice (2018) argues that the explanatory contribution of "minimal" (i.e. highly idealized) models in for example physics and biology are best seen in terms of providing counterfactual information about real world targets, and that the models' "holistic" distortion of their targets contributes to this function.

Modal modeling can also be *hypothetical* or *exploratory*. Such modeling is important in situations where a putative target lacks theoretical description consisting of shared and widely accepted principles and concepts. The possibility claims allegedly established by such models can be used for a number of different purposes, such as providing proofs of principle (Gelfert, 2016), refute necessity or impossibility claims (Grüne-Yanoff, 2013), or delineate the space of what is possible and impossible regarding a putative or actual phenomenon (Massimi, 2019). Such modeling may also provide "explanations in search of observations" (Sugden, 2011), that is offer representations of possible properties of possible explananda, for the purpose of understanding such phenomena in case they become actual. Most of the models mentioned above are abstract and mathematical in nature, but there are examples of more concrete modal modeling too. For instance in synthetic biology, where nonactual possibilities are studied with the help of concrete models used to represent minimal cells and alternative genetic systems even though such targets might turn out to be only partially (if at all) realizable in practice (Knuuttila & Koskinen, 2020; Koskinen, 2017).

Finally, while the clearest examples of modal modeling tend to deal in possibility claims or counterfactual conditionals, not all do. Nguyen (2020) denies that highly simplified "toy models" common in biology and economics only support "mere" how-possibly explanations. Rather, facts about toy models can sometimes be translated into claims which ascribe properties–for example *capacities* or *susceptibilities*–to real world targets. For example, in Akerlof (1970) "market for lemons" model, asymmetric information prevents car trades from occurring, despite the fact that at any given price there are sellers willing to sell their car and buyers willing to buy it. When properly interpreted in terms of a given target, Nguyen suggests, the claim supported by the model is something like: an asymmetric information state in this (particular, real world) market *increases the market's susceptibility to fail* to reach Pareto-efficient equilibrium. Importantly, this still makes Akerlof's model an instance of modal modeling, supporting ascription of a *de re* modal property–a susceptibility, which is a form of disposition–to a particular system.

3 | THE EPISTEMIC QUESTION FOR MODAL MODELING

It is a familiar fact that models, insofar as they are to contribute to science's uncovering of information about the world, must satisfy some "standards of accuracy" (Frigg & Nguyen, 2016) or be "faithful epistemic representations" (Contessa, 2007). Such standards are claimed to ensure the models's epistemic function: they underwrite—appropriately circumscribed—inferences from these models to new knowledge about the world. Now, modal modeling accounts—especially of practices that are not overtly modal but nevertheless are interpreted as such—are often motivated by the fact that the models in question apparently do not satisfy standard criteria of representational accuracy (e.g. model-target similarity, or isolation of relevant causal factors), given by more traditional accounts of how we learn from scientific models. This prompts the question: if this model does not represent an actual target with accuracy, what kind of epistemic function could it play? Modal modeling accounts suggest that such models can provide modal knowledge, and that this is their (immediate) epistemic contribution.

This claim prompts several questions. One concerns accuracy in description of the role(s) that these models actually play in scientific practice: does the modal interpretation really fit with the observed practices? Another one concerns whether the modal modeling account does better than standard accounts in explaining the epistemic import of these modeling practices: the model does not the satisfy the criteria associated with successful non-modal modeling, but does it satisfy the criteria associated with successful modal modeling?

That in turn raises the deeper questions of under what conditions, and in virtue of what, scientific models can plausibly perform the epistemic function of supporting conclusions about what is possible, or what would be the case under counterfactual circumstances. To be sure, modal modeling might not need to satisfy standard criteria associated with non-modal modeling, but this does not mean that the models do not need to satisfy *any* standards. The onus is on the modal modeling accounts to say something about the conditions under which a model is a reliable scientific tool for eliciting modal information–whether about possibilities, dispositions, or counterfactual conditionals.

We assume that with modal claims, just as with non-modal claims, there is generally a fact of the matter as to whether they are true or false—the truth-value of modal claims are not matters of opinion. When claiming that such-and-such is possible, one can be right, or one can be wrong. In order to be justified in claiming that such-and-such is possible, one needs to give reasons, for example by citing relevant evidence or following the right inference procedures. Modal modeling accounts state that appealing to scientific models is a legitimate way of justifying a modal claim. This places certain constraints on the models, since not just any model can be trusted to relay reliable modal information.

Take Thomas Schelling's much-discussed checkerboard model, mentioned above. It consists of a two-dimensional grid with two types of individuals, say triangular and circular, initially distributed randomly on the grid. All individuals move iteratively according to one rule: if not at least one-third of one's neighbors are of one's own type, move to an empty square where that condition is fulfilled. Schelling interpreted this rule as individuals being happy to "live" in mixed areas, as long as they are not in a strong minority (1971, 148). After relatively few iterations, this generates a "segregated" grid with triangles and circles in separate clusters. The thus interpreted model is sometimes claimed to describe a *possible* cause of segregation: it is possible that segregation occurs in a population as the result of an individual preference to not be in strong minority. Assuming that there is a fact of the matter regarding whether this mechanism *is* a possible sufficient cause of segregation or not, and that Schelling was justified in claiming, with reference to the model, that it is, there must be something about the checkerboard model in virtue of which it *does* provide good reasons for this claim. What is that something, and why think it makes the model a good guide to a modal truth? More generally, what must models or modeling practices be like in order to provide good reasons for modal claims (*under what conditions*), and why does fulfilling those criteria make models epistemically valuable with respect to modal truth (*in virtue of what*)? This is the two-part epistemic question for modal modeling. The two parts are closely intertwined in the sense that answering the latter is part and parcel of a successful defense and motivation of an account in response to the former.

To be clear, the modal modeling claim is *not* that modeling is the main, let alone the only, way for scientists to come by modal knowledge. Rather, the claim is that modeling is *one* way for scientists to arrive at justified modal claims. The epistemic question challenges philosophers of science to elucidate why we should accept that this is so, concentrating on the *justificatory* part rather than the alleged fact that scientists *in fact* make modal claims with reference to models.

4 | THREE ATTEMPTS TO ADDRESS THE EPISTEMIC QUESTION

The epistemic question has so far received too little attention in the literature. But there are some significant exceptions where, in addition to documenting and analyzing modal modeling practices, philosophers have also attempted to sketch an answer to the epistemic question, and in particular its 'under what conditions' part. We survey them in this section. But first, note that these accounts should not be seen as mutual exclusives: different models may support justified modal conclusions in virtue of different kinds of properties. In fact, since modal information of different kinds (e.g. epistemic, objective) are relevant to different modeling practices, we should expect different answers to the epistemic question to be correct for different modal models. Given the great diversity of models and modal claims at play then, it seems wise to approach the topic in a pluralist spirit.

4.1 | The universality account

Robert Batterman and Collin Rice have argued that a certain class of models ("minimal modals") are best interpreted as serving the epistemic function of drawing true counterfactual conclusions about real world targets, and thus contribute to or even provide how-actual explanations through modal information. Granting this modal modeling claim, however, raises the question of

> how holistically distorted models can provide *true* counterfactual information about their target systems (…) I will try to offer one possible solution to this problem by appealing to *universality*." (Rice, 2018, 2812)

That is, Rice undertakes to answer the epistemic question as it arises for the modal interpretation of minimal models, in terms of universality. Basically, the idea is that if a model system and a target system are *in the same universality class*, then one can justifiably use the model in order to learn about what would happen to the target system under such-and-such circumstances. Being in the same universality class amounts to being disposed to exhibit the same macrobehavior. Importantly, systems that are in the same universality class can be very different in terms of what properties or physical components they have–certain macrobehaviors of systems in the world are "largely independent of the details of their physical components" (Rice, 2018, 2812). This means that we can learn from the model about the target's counterfactual behavior despite the model having very different properties or structure, and it means that the same idealized model can sometimes be used to elicit counterfactual information about several, very different, target systems. They give some examples of this, including the Lattice Gas Automaton model vs. real fluid flow, Fisher's linear substitution cost model vs. sex ratios in animal populations (both in Batterman & Rice, 2014), and an optimal foraging model in an infinite population vs. Eider duck foraging behavior (Rice, 2018).

Whether some systems are in the same universality class is an empirical question, something that scientists need to discover (Rice, 2018, 2813; 2019, 200). Clearly, one cannot come to know this by looking at the microproperties or components of systems, since the very point about universality is that some behaviors are independent of those factors. Batterman and Rice (2014) discuss in detail how stability of behaviors across various changes have been established for the Lattice Gas Automaton (LGA) model. This example relies on the renormalization group strategy, illustrated by the Kadanoff block spin transformation. The purpose of such investigations is to identify the physical systems, consisting of many interacting entities, that share the same scale-invariant macrobehavior, for example transitioning to an orderly state below a certain transition temperature. The basic idea of the Kadanoff transformation is to start with a space of possible systems, in this case constrained by assumptions about its entities and interactions. Some members of this possibility space might be real fluids, others include the LGA of a certain dimensionality and scale. Each of them exhibits some macrobehavior, and some of them might exhibit the same. A rescaling procedure applied to each possible system aggregates entities into a group and determines the macrobehavior of this rescaled system. Systems that continue to exhibit the same macrobehavior under such transformations are identified to belong to the same universality group, thus justifying the use of LGA to explain the behavior of fluids belonging to the same universality class. The general idea—that by performing transformations on the set of models, and showing that their behavior converges despite these changes, one justifies inferences to *counterfactual independence*, that is one comes to know about changes to the system that do not make a difference to the observed behaviors—is presumably the same across other attempts to establish universality classes, but it remains an open question how the process is to be constrained in cases where the renormalization strategy is not relevant.

4.2 | The credibility account

Several authors have suggested that simple, highly idealized toy models—common in economics and biology—provide knowledge of possibility (e.g. Grüne-Yanoff, 2009; Reutlinger et al., 2018; Weisberg, 2013). But in virtue of what? Robert Sugden (2000) has famously suggested a response to the epistemic question for this modal modeling account in terms of credibility (taken on board by e.g. Grüne-Yanoff, 2009; Mäki, 2009, 39-40; Morgan, 2012; Betz, 2015). The idea is that toy models can serve this epistemic function insofar as they are *credible*. The credibility of a model, Sugden argues, contributes to the justification of inference from it to the real world: "we can have more confidence in them [the inferences], the greater the extent to which we can understand the relevant model as a description of how the world could be" (2000, 24).

This credibility thesis immediately prompts the question of what it means for a model to be credible, and Sugden unfortunately remains ambiguous in his answer, saying that a model is credible if the situation it depicts "could be real", or is "parallel to the real world", in the sense that it conforms to our experiences and intuitions about the causal forces that operate in the real world (2000, 25). The most developed attempt to elucidate Sugden's notion of credibility is due to Grüne-Yanoff (2009). He picks up on an analogy with fiction that Sugden suggests when writing that credibility in models is "rather like credibility in 'realistic' novels" (2000, 25). This emphasizes the role of the imagination, which plays a central role in most accounts of fiction. When we read a fictional text, we imagine a fictional world, proceeding from the text but going beyond it by adding detail, drawing out implications and filling in gaps. When we judge a fiction to be credible, it is this imagined world that we are judging. Analogously, the formal structures of toy models are interpreted by their users as imaginary worlds or scenarios, proceeding from but going beyond the explicit model description. This imagined model world is then assessed for credibility. If it is credible, that is a reason to think that the model result (or some equivalent thereof) is possible. For instance, Schelling's checkerboard model provides us with a reason to think that it is possible for racial segregation to result despite individual citizens' having a preference for living in mixed areas (or differently put, that a mere preference for not being in strong minority is a possible mechanism behind segregation), *because the checkerboard model describes a credible scenario*.

But in virtue of what is the imagined world that Schelling's model prompts credible? More generally, what must a model world be like in order to be credible? Grüne-Yanoff stresses that many particular features of a fictional world can deviate extensively from what the actual world is like, yet it can be judged credible–the same goes for model worlds. But it must be internally coherent: sufficiently detailed and free of incoherent or contradictory assumptions and implications. Moreover, development in the imagined world—what "happens" in the fiction/model—must be judged to be plausible *conditional on* the background information provided about for example preferences, environment, and so on. These conditional judgements are, Grüne-Yanoff writes, "driven by empathy, understanding, and intuition" (2009, 94-95). Presumably, not all users of a model are in a good position to reliably judge whether the model world *does* fulfill this second condition. As the mention of "understanding" indicates, only the assessment of a *competent user* of the model will do—someone with the appropriate background knowledge and experience, that is. To sum up, a model is credible—and hence a good guide to justified possibility claims—when the model world is internally coherent and a competent user of the model would judge the development in the model world to be intuitively plausible, conditional on the model setup. Notably, and in contrast to the conceivability account discussed in the next section (as well as Sugden's account, which proposes both the fiction analogy *and* the laws of nature constraint), this version of the credibility account does not rely on being constrained by laws of nature.

4.3 | The physical conceivability account

In discussing the epistemic contribution of a class of models that she, following Gelfert (2016), calls "exploratory", Michela Massimi suggests that they deliver "genuine modal knowledge (about how things *might be in nature*)" (2019, 871), in particular two kinds of possibility knowledge. "Hypothetical" exploratory modeling provide knowledge what is "objectively possible", for example about whether a hypothetical target can exist, and "fictional" exploratory modeling provides knowledge of what is "causally possible" for actual targets. She provides one example of each: modeling of SUSY (supersymmetrical) particles in physics, used establish possible ways in the theorized particle could be, and Maxwell's honeycomb model of the ether used to "identify a possible causal mechanism behind" (2019, 872) the actual phenomenon of electromagnetic induction.

Massimi gives an answer to the epistemic question for modal modeling in terms of a form of constrained imagination that she calls "physical conceivability". More generally speaking, the key claim is that if a scientist can physically conceive of *p*, she is justified in believing that *p* is possible. The notion of 'conceiving' as a particular mode of imagining is a familiar one from the literature on the epistemology of modality—we touch on this in section 5 below—and the physical conceivability account is similar to the credibility accounts above in appealing to imagination, but there are important differences. Massimi's physical conceivability account differs, from both traditional conceivability theory and from the credibility account above, in assigning a large role to (purported knowledge of) the laws of nature. In particular, physical conceivability explicitly requires that the imagining is constrained by the imagining subject's knowledge of the laws of nature:

> *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)* (Massimi, 2019, 872, our emphasis).

How this works is especially clear in the SUSY example. As an important step towards confirming whether there are any Salis, scientists investigate the different ways in which it is physically possible that a SUSY particle exists. For this purpose, scientists have devised a modeling technique–the pMSSM-19 – which produces different model scenarios, each of which portrays the SUSY particle as having different properties and value assignments (e.g. a given mass value, a given decay mode) for 19 parameters, and as consistent with certain nomological constraints. The idea is presumably that the scenarios produced by the pMSSM-19 are physically conceivable, and so for each such model scenario, it is concluded that a SUSY particle, as conceived in the scenario, is physically possible.

The case of Maxwell's ether model is supposed to work slightly differently, and to deliver a different kind of modal claim, that is a possible causal mechanism behind an actual phenomenon. Just how it is supposed to work, and why it licences a different type of modal knowledge, remains somewhat underdeveloped in Massimi's paper. The key claim appears to be that the interplay between imagining and knowledge of laws is different: the conceiving is now not "law-bounded" as in the SUSY case, but "law-driven". The connection between fictional scenario (the ether system) and specific actual phenomenon (electromagnetic induction) is allegedly established via analogical reasoning with models in other domains and the laws they incorporate (in this case Helmholtz's equations for fluid dynamics). Whether such an inference is justified obviously depends on the justification of the employed analogy, but Massimi remains vague about the nature of this justification; in particular, it remains unclear in what sense laws and physical conceivability can "drive" analogy in a way that would be epistemically relevant to justification of the type of modal claim at issue.

5 | JUSTIFICATORY STRATEGIES

The surveyed accounts sketch different answers to the question of under what conditions models provides certain modal information (e.g. when they are credible, physically conceivable, or in a certain universality class) but they say little of why these conditions make models good guides to true modal information (the 'in virtue of' part of the epistemic question). In this section and the next we try to lay bare the rough justificatory strategies that the accounts rely on, and raise some important questions that each would need to address in order to have a full answer to the epistemic question for modal modeling.

The credibility account and the physical conceivability account have in common that they both take the *imagination* to be centrally involved in modal modeling. This lines up well with a more general view of modeling as centrally involving the imagination, most clearly represented by the *Waltonian fictionalist* account of models (Frigg, 2010; Levy, 2015; Salis, 2021; Toon, 2012), according to which all models are games of make-believe that scientists collectively engage in. But it also taps into a venerable tradition in the epistemology of modality, going back to Descartes and Hume, according to which knowledge of possibility and/or counterfactual conditionals is acquired through the imagination (e.g. Yablo, 1993; Kung, 2010; Williamson, 2007). In the modal epistemology debate, it has long been discussed how imagination is to be properly constrained in order to be a reliable guide to modal truth, in light of the fact that we can clearly imagine things that are impossible, or imaginatively develop a counterfactual scenario in a deviant way. One view that has emerged from this is discussion is that although imagination may often be involved in modalizing, the justificatory work is really done by some constraining background knowledge (tacit or explicit). For instance, one is justified in believing that *p* is possible on basis of being able to imagine a world in which *p*, just in case one's imagining was properly constrained by certain accurate background assumptions.

That imagination needs to be constrained in order to serve science has not gone unnoticed, neither in general (see e.g. Skolnick Weisberg 2020; Salis & Frigg, 2020) nor by modal modeling theorists. For instance, Sugden appears to consider knowledge of actual laws of nature as restrictions on what imagined worlds count as credible (2000, 25);

Massimi's law-bounded physical conceivability is explicitly constrained by such knowledge; and the requirement that reliable credibility judgements need to be issued by competent users of the model also echo this concern that not just any imagined scenario will do. This suggests that whatever the justificatory role of imagination in (some) modal modeling might be, at least some of the justificatory power derives from some background knowledge or other.

The universality account is different from the two others, in that it does not appeal to imagination or conceiving. However, it too apparently relies on substantial background knowledge. Being justified in relying on a minimal model for counterfactual information about the target requires knowing, or being able to reliably judge, that the model and target are disposed to behave in the same way. The authors outline a few different cases in order to illustrate how that might be established, but none of them are as detailed or thorough as the case with the LGA model, and as a result for example the Eider foraging behavior example comes across as less plausible. This clearly brings out the fact that the processes of establishing universality classes are reliable only if constrained or supported by a fair amount of background knowledge, and that it is a nontrivial issue whether we possess it in a given case.

Moreover, it is plausible that systems in one universality class are not disposed to behave in the same way in all respects. Instead, universality class membership is with respect to some macrobehavioral similarity or other. And as Rice notes, "the universality class required to justify a particular instance of idealized modeling will depend on the details of the modeling context" (2018, 2816). That is, what behavioral similarity is relevant depends on what modal information one is after. But this raises the question of how we come by the knowledge, or ability to reliably judge, what behavioral similarities or dissimilarities are relevant to the truth of a prospective counterfactual claim. In a sense, this is just an instance of a more general issue that besets any modeling account that appeals to similarity (whether of features or behavior) in their standards of accuracy: what similarities are relevant? It need not be a problem that this must be specified (Giere 1999, 2010). But doing so requires a fair amount of background knowledge about both model and target. As with the modal modeling accounts that appeal to imagination, there may be interesting connections to modal epistemology, see for example Roca-Royes (2017) for a similarity-based theory of modal justification, which Sjölin Wirling (forthcoming) argues can be extended with scientific models.

In sum, while ability to imagine, and established behavioral similarity, respectively are key to the justificatory strategies employed here in response to the epistemic question for modal modeling, background knowledge of various kinds turn out to be very important too, on closer inspection.

6 | QUESTIONS FOR FUTURE RESEARCH

The fact that background knowledge appears to be doing a lot of the justificatory work presents modal modeling with a number of issues that have not been addressed in the extant literature but which deserve careful consideration. We identify three such issues here.

First, what background knowledge is required to drive, constrain, or assess (the product of) modal modeling partly depends on what kind of modal information one is after: is it epistemic or objective possibility, and if the latter, is it physical, biological, political, or economical possibility? Not all possibilities are of interest in science, and different background knowledge will be relevant to different senses of 'possible'. This is not always as clearly distinguished in the literature as it needs to be. For instance, while knowledge of laws of nature may place suitable constraints on modeling in particle physics given for example the aim of finding out what physically possible ways there are for a SUSY particle to exist in this world, it is not at all clear that knowledge of the laws of nature are sufficient, or even relevant, to constrain modeling in for example biology or economics (Grüne-Yanoff, 2009, 92; Sjölin Wirling, 2021). Moreover, when it comes to possibility claims, what background knowledge needs to be brought to bear on a modeling exercise depends on whether the possibility is epistemic or objective (Grüne-Yanoff & Verreault-Julien, 2021; Sjölin Wirling & Grüne-Yanoff, forthcoming). It is an important task for modal modeling accounts to specify the type of possibility of interest *and* elucidate the background knowledge relevant to it.

Second, if background knowledge is doing the justificatory work in modal modeling, this raises a puzzle. Modal modeling is often allegedly exploratory, crucial to advancing science in areas where there is a *lack* of confirmed theory or shared background knowledge. It is also claimed that modal modeling helps *challenge* the background assumptions in a field, by showing that so-and-so is possible despite it contradicting some widely held general principle or law. Finally, models that are allegedly modal (e.g. toy models) are often taken to work *independently of framework theory* (cf. Reutlinger et al., 2018) There is a very delicate balance to be struck between using modal modeling for those purposes, and satisfactorily answering the epistemic question for modal modeling.

Third, if modal modeling centrally involves imagination, as on the credibility and physical conceivability accounts, there is the question of what, more exactly, the justificatory role (if any) of imagination is. Does the imagining itself contribute anything, or is a "mere" heuristic, exploratory tool? How does it interplay with background knowledge in order to deliver modal information that we have reason to trust?

In sum, this brief investigation of modal modeling accounts has revealed new avenues of research that we believe deserve further scrutiny, and there are surely more to be found on even closer inspection.

ACKNOWLEDGEMENT

This research was conducted as part of the project *Towards a Methodology of Modal Modeling*, funded by the Swedish Research Council grant no. 2018-01353. Sjölin Wirling also acknowledges funding from the Swedish Research Council grant no. 2019-00635.

ORCID

Ylwa Sjölin Wirling https://orcid.org/0000-0002-5638-4192

ENDNOTE

[1] To be clear, some of the possibilities (both objective and epistemic) probed in modal modeling may well in fact be actualized (although not known to be). The point here is that modal models do not, or are not expected to, deliver information about what is actually the case with some target, but modal information in some form or other.

REFERENCES

Akerlof, G. A. (1970). The market for "Lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*(3), 488–500.

Batterman, R. W., & Rice, C. C. (2014). Minimal model explanations. *Philosophy of Science*, *81*(3), 349–376.

Betz, G. (2015). Are climate models credible worlds? Prospects and limitations of possibilistic climate prediction. *European Journal for Philosophy of Science*, *5*(2), 191–215.

Bokulich, A. (2014). How the tiger bush got its stipes: 'How Possibly' vs. 'How Actually' model explanation. *The Monist*, *97*(3), 321–338.

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

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

Frigg, R., & Nguyen, J. (2016). Scientific Representation. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2020 edition.

Gelfert, A. (2016). *How to do Science with Models: A Philosophical Primer*. Spring International.

Gelfert, A. (2019). Probing possibilities: Toy models, minimal models, and exploratory models. In M. Fontaine, C. Barés-Gómez, F. Salguero-Lamillar, L. Magnani, & Á. Nepomuceno-Fernández (Eds.), *Model-Based Reasoning in Science and Technology* (pp. 3–19). Springer International.

Giere, R. N. (1999). Using models to represent reality. In L. Magnani, N. J. Nersessian, & P. Thagard (Eds.), *Model-Based Reasoning in Scientific Discovery* (pp. 41–57). Kluwer Academic/Plenum Publishers.

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

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

Grüne-Yanoff, T. (2013). Appraising Models Nonrepresentationally. *Philosophy of Science*, *80*(5), 850–861.

Grüne-Yanoff, T., & Verreault-Julien, P. (2021). How-possibly explanations in economics: anything goes? *Journal of Economic Methodology*, *28*(1), 114–123.

Knuuttila, T., & Koskinen, R. (2020). Synthetic fictions: turning imagined biological systems into concrete ones. *Synthese*. https://doi.org/10.1007/s11229-020-02567-6
Koskinen, R. (2017). Synthetic biology and the search for alternative genetic systems: Taking how-possibly models seriously. *European Journal for the Philosophy of Science*, *7*(3), 493–506.
Kung, P. (2010). Imagining as a Guide to Possibility. *Philosophy and Phenomenological Research*, *81*(3), 620–663.
Levy, A. (2015). Modeling without models. *Philosophical Studies*, *172*(3), 781–798.
Mäki, U. (2009). MISSing the world. models as isolations and credible surrogate systems. *Erkenntnis*, *70*(1), 29–43.
Massimi, M. (2019). Two kinds of exploratory models. *Philosophy of Science*, *86*(5), 869–881.
Morgan, M. (2012). *The World in the Model: How Economists Work and Think*. Cambridge University Press.
Nguyen, J. (2020). It's not a game: Accurate representation with toy models. The *British Journal for the Philosophy of Science*, *71*(3), 1013–1041.
Reutlinger, A., Hangleiter, D., & Hartmann, S. (2018). Understanding (with) toy models. *British Journal for the Philosophy of Science*, *69*(4), 1069–1099.
Rice, C. (2018). Idealized models, holistic distortions, and universality. *Synthese*, *195*(6), 2795–2819.
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). Springer International.
Rohwer, Y., & Rice, C. (2013). Hypothetical Pattern Idealization and Explanatory Models. *Philosophy of Science*, *80*(3), 334–355.
Salis, F. (2021). The New Fiction View of Models. *British Journal for the Philosophy of Science*, *72*(3), 717–742.
Salis, F., & Frigg, R. (2020). Capturing the scientific imagination. In P. Godfrey-Smith, & A. Levy (Eds.), *The Scientific Imagination* (pp. 17–50). Oxford University Press.
Schelling, T. C. (1971). Dynamic Models of Segregation. *Journal of Mathematical Sociology*, *1*(2), 143–186.
Sjölin Wirling, Y. (forthcoming). Extending similarity-based epistemology of modality with models. Ergo: An Open Access Journal of Philosophy.
Sjölin Wirling, Y., & Grüne-Yanoff, T. (fortcoming). Epistemic and Objective Possibility in Science. *British Journal for the Philosophy of science*, https://doi-org.ezproxy.ub.gu.se/10.1086/716925
Sjölin Wirling, Y. (2021). Is credibility a guide to possibility? A challenge for toy models in science. *Analysis*.
Skolnick Weisberg, D. (2020). Is imagination constrained enough for science? In P. Godfrey-Smith & A. Levy (Eds.), *The Scientific Imagination* (pp. 250–261). Oxford University Press.
Smith, J. M., & Price, G. R. (1973). The logic of animal conflict. *Nature*, *246*(5427), 15–18.
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 and Philosophy*, *26*, 717–736.
Toon, A. (2012). *Models as Make-Believe: Imagination, Fiction, and Scientific Representation*. Palgrave-Macmillan.
Verreault-Julien, P. (2017). Non-causal understanding with economic models: the case of general equilibrium. *Journal of Economic Methodology*, *24*(3), 297–317.
Verreault-Julien, P. (2019). How could models possibly provide how-possibly explanations? *Studies In History and Philosophy of Science Part A*, *73*, 22–33.
Weisberg, M. (2013). *Simulation and Similiarity: Using Models to Understand the World*. Oxford University Press.
Williamson, T. (2007). *The Philosophy of Philosophy*. Blackwell.
Yablo, S. (1993). Is Conceivability a Guide to Possibility? *Philosophy and Phenomenological Research*, *53*(1), 1–42.
Ylikoski, P., & Aydinonat, N. E. (2014). Understanding with theoretical models. *Journal of Economic Methodology*, *21*(1), 19–36.

AUTHOR BIOGRAPHIES

Ylwa Sjölin Wirling is a Swedish Research Council postdoctoral researcher at the University of Gothenburg and a visiting researcher at the University of Manchester. She was previously a postdoc at the KTH Royal Institute of Technology in Stockholm, and received her PhD from the University of Gothenburg in 2019. She works in epistemology (of modality, and more generally), metaphysics, and the philosophy of science. Her previous work at the intersection between modal epistemology and scientific modeling has been published in journals like *Analysis* and *Ergo*.

Till Grüne-Yanoff is a professor of philosophy at KTH Royal Institute of Technology in Stockholm. He specializes in philosophy of science, philosophy of economics, and decision theory. He received his PhD in 2004 from the London School of Economics and Political Science (LSE), and has previously held positions at University of Helsinki and the Helsinki Collegium of Advanced Studies. He is the editor of *Economics & Philosophy*, and currently directs

two research projects; one on modal modeling in science, and one on the mechanisms underlying behavioral policies for sustainable energy consumption.

How to cite this article: Sjölin Wirling, Y., & Grüne-Yanoff, T. (2021). The epistemology of modal modeling. *Philosophy Compass*, *16*(10), e12775. https://doi.org/10.1111/phc3.12775