Synthese (2023) 201:161
https://doi.org/10.1007/s11229-023-04135-0

ORIGINAL RESEARCH

Modals model models: scientific modeling and counterfactual reasoning

Daniel Dohrn[1]

Received: 28 December 2021 / Accepted: 18 March 2023

Abstract
Counterfactual reasoning has been used to account for many aspects of scientific reasoning. More recently, it has also been used to account for the scientific practice of modeling. Truth in a model is truth in a situation considered as counterfactual. When we reason with models, we reason with counterfactuals. Focusing on selected models like Bohr's atom model or models of population dynamics, I present an account of how the imaginative development of a counterfactual supposition leads us from reality to interesting model assumptions; how it guides our reasoning from these assumptions to interesting consequences for the model scenario via counterfactual entailment; and how it leads us back to conclusions on real target phenomena.

1 Preliminaries

1.1 The Prima facie case for using counterfactuals

Philosophers have claimed that *counterfactual conditionals* or briefly counterfactuals, usually expressed as *if A were/had been the case, C would/would have been the case* (*A*□→*C*), play an important role in science, for instance in formulating laws and explanations (e.g. Goodman, 1947; Lewis, 1986; Woodward, 2003). More recently, some philosophers have argued that they also play a role in *scientific modeling* (e.g. Giere, 1988; Adams, 1993; Bokulich, 2011; Psillos, 2011; Williamson, 2017, 2020; Tan, 2019; Godfrey-Smith, 2020; McLoone, 2021).

✉ Daniel Dohrn
Daniel_Dohrn@yahoo.com

1 Dipartimento di Filosofia, Università degli Studi di Milano, Via Festa del Perdono 7, 20122 Milano, Italy

Published online: 26 April 2023

I shall say some words on why to use counterfactuals in modeling. Normally model descriptions are not taken to be literally true of anything real.[1] One may say that they are simply false, but this makes it the more puzzling how they could be of any avail to science (van Riel, 2015). One salient alternative is that there is something model descriptions should be true of. The question becomes *what* they are true of. I shall take a quick look at some candidate answers.

One tempting answer is that models are fictions (e.g. Frigg, 2010; Toon, 2012; Salis & Frigg, 2020; Thomasson, 2020). Truth in a model can be defined by truth in fiction. Yet such an answer faces serious difficulties. Although fictions and make-believe usually integrate aspects of reality, there is no systematic transition from what is the case in a fiction to reality (see Levy, 2015, 789–790). Counterfactuals seem to fare better in this respect:

> 'Without the possibility of discharging suppositions and speaking from outside them, scientific statements cannot be held to account against physical reality… By contrast, no such move is normally available *within* a game of make-believe or a fiction.' (Williamson, 2020).

Transitions to reality are built into the very logics of counterfactuals, e.g. by reasoning: $A \Box\!\!\rightarrow C$*; A; thus C.* As we will see, the transition to reality becomes more complicated in the case of models, but at least there is a strong systematic relationship between the truth of a counterfactual and what is true in reality.

To be sure, fictions may be analyzed by counterfactuals (Lewis, 1983) or vice versa. Yet it is not obvious what adding the notion of a fiction positively contributes over and above a direct counterfactual-based account. Such an account can even take on board a view of models as particular (imagined or real) objects that act as props for a game of make believe (Godfrey-Smith, 2006, 732; Frigg, 2010, building on Walton, 1990). Such props may as well be read as initiating a supposition: if things were like that/as imagined…?[2]

I don't deny that there are further resources for spelling out the contribution of fiction to scientific explanation and modeling. The huge literature on artifacts and on philosophical fictionalism may harbour such resources. Thus, my arguments provide only a prima facie case for a counterfactual-based alternative.

An alternative answer is that models are true of *possible* situations. Possibility can be interpreted in different ways. It is standard to distinguish between an objective (circumstantial, alethic) and an epistemic sense of possibility. As for epistemic possibilities, we often rule out the model assumptions as hypotheses on what the actual world is like. Then the model situation cannot simply be interpreted as a relevant epistemic possibility. It may be interpreted as an *a priori* possibility, but the question becomes in how far such a possibility is relevant to empirically informed science.

[1] Mutatis mutandis my considerations apply to non-descriptive representations like pictures.

[2] This holds also for Levy's (2015, 4.1.) proposal that the props are the target phenomena, albeit under a fictional model description. Levy's example is an ideal gas model: 'The model is an instruction to regard (imagine) a real world gas as if it had various features (including non-colliding molecules).' (Levy, 2015, 794) The idea can perfectly be represented by a counterfactual that invites to imaginatively develop the supposition of a real gas having various features.

The remaining alternative are objective possibilities. In this case, again Williamson's point about fictions applies. We need to reason from merely possible situations to reality. Again counterfactuals provide the most eligible transition from merely possible situations to reality.

A further problem that is specific to treating model situations as possible is that many typical models do not seem to describe possible situations, possibility taken in the widest objective sense that is usually called 'metaphysical'. For instance, modeling predators and prey as continuous in number as in the Lotka-Volterra model typically does not involve a commitment to the metaphysical possibility that they come in continuous quantities. One could think of possibility in an even wider sense, e.g. conceptual or logical possibilities, but the less restrictive the relevant sense of possibility becomes, the more difficult the transition to reality will be.[3]

Again, there is the alternative of using counterfactuals. It is not trivial that construing model situations by counterfactuals avoids the problem, but it opens up further room for manoeuvre. In particular, there is a thriving literature on *non-trivial counterpossibles*, which raises prospects for dealing with impossible model situations (Tan, 2019; McLoone, 2021).

I mention a further alternative: one may construe models as abstract objects (discussion in Levy, 2015, Sect. 2). Truth in a model may be defined as truth with regard to such an object. We may need abstract objects to play a role in scientific explanations anyway, for instance in understanding the contribution of mathematics, and a theory of models may fall out of an account of such a contribution. The alternative so far is too sketchy to weigh its pros and cons, so I content myself with acknowledging its relevance. I mention just one minor point: there are counterfactual-based approaches especially to the role of mathematics in scientific explanation (Baron et al., 2017). Rather than competing with a construal of models as abstract objects, a counterfactual-based approach may provide an attractive framework for dealing with the role of such objects in scientific reasoning.

In sum, there are *prima facie* reasons for taking a counterfactual approach to models seriously. I shall now develop my own proposal for such an approach, introducing a further key notion, *imagination*, which I interpret as involving a capacity of finding and developing counterfactual suppositions.

1.2 Limiting the scope: paradigmatic models

I shall clarify what kinds of models I intend to cover. While I remain neutral about a general understanding of science in terms of models (Suppes, 1960), my understanding of modeling draws on Weisberg : 'modeling is just one kind, albeit an important kind, of theorizing… Modeling… is the *indirect* theoretical investigation of a real world phenomenon using a model.' (2007, 209, m.e.) I take the direct/indirect distinction from Weisberg. Modeling represents indirectly, as contrasted to direct, albeit abstract representations of reality as in Mendeleev's periodic system. The distinction

[3] One may doubt that there is a conceptual possibility of e.g. rabbits coming in continuous numbers anyway.

can be accounted for by saying that model descriptions are true in counterfactual situations. In contrast, the periodic system describes the actual world.

Models are used for very different purposes in science. Many categories of models are distinguished (Frigg & Hartmann, 2020). It would be overambitious to cover them all. Instead, I shall pick up some examples which have been used by proponents of a counterfactual-based approach:

Bohr's atom model (Bohr, 1913; Bokulich, 2011, 41–42);
A model of the solar system with point masses (Adams, 1993, 5; Tan, 2019, 45);
A model of water as a continuous, incompressible medium (Tan, 2019, 47);
A model of an ideal, massless-string pendulum (Tan, 2019, 44; Williamson, 2020);
Models of population dynamics, e.g. Lotka-Volterra (Jenkins & Nolan, 2012; McLoone, 2021).

My approach aims at capturing the role of counterfactual suppositions in models *like these*. I surmise that the lesson can be *mutatis mutandis* extended to other models, but I shall leave such an endeavour to future debate. One may doubt that the models mentioned form a unified category, but the only commonality I commit myself to is that my counterfactual-based approach applies to them. The models mentioned all seem to involve impossible assumptions, but my considerations are not limited to such assumptions.[4]

I shall distinguish model *assumptions* from model *descriptions*. Model assumptions are the counterfactual suppositions which determine the content of the model. A model description captures what is true in a model (Giere, 1988, 79). Of course, explicit model descriptions will often be merely partial descriptions of what is true in a model. A model description is true of the relevant counterfactual situations, for instance the closest possible worlds in which the model assumptions are true.

2 Counterfactual imagination in scientific models

I again follow Weisberg in identifying 'three stages' of modeling:

> 'In the first stage, a theorist constructs a model. In the second, she analyzes, refines, and further articulates the properties and dynamics of the model. Finally, in the third stage, she assesses the relationship between the model and the world if such an assessment is appropriate.' (2007, 209).

I do not interpret the three stages as forming a chronological order. I give my own account of them, which diverges from Weisberg's: many models are not constructed

[4] Point masses, water as a continuous, incompressible medium, massless-string pendula, and animal populations that are continuous in number seem metaphysically impossible. Bohr's model presumably is nomically impossible due to violating Heisenberg's uncertainty principle. It may be metaphysically impossible if we assume the laws of nature to be metaphysically necessary.

out of the blue; they take a conspicuous departure from reality. In turn, if models have a role in accounting for real target phenomena, there should also be a way back to reality. Following a pathway laid out by Godfrey-Smith (2020), I shall explore the role of counterfactual imagination with regard to both stages. Following Weisberg, I add a third stage: the imaginative development of the counterfactual supposition towards further features of the model situation. My main innovation in this paper is a differentiated account of the role of counterfactuals with regard to all three stages (Sect. 2.1.-2.3.).

2.1 From reality to models

In this section, I focus on the departure from reality. My discussion advances debate in assigning imagination a nuanced role: one of several roles of imagination is to creatively explore practically relevant alternatives to reality by simulating them. The simulation is streamlined by explicit and implicit information on the actual world (Sect. 2.1.1.). I draw a new parallel between practically relevant and theoretically interesting alternatives (2.1.2.).

2.1.1 Imagination between Creativity and Restraint

Godfrey-Smith emphasizes the practical relevance of our capacity of creatively exploring salient alternatives to reality and its connection to conditional thinking:

> ‘Plausibly, the idea of possibility has a primitive association with action: the world at large determines how things *are*; we determine what to *do*, and in these episodes we take ourselves to choose from possibilities. From there, a sense of possibility projects backward and sideways. We see other events, including past events, as embedded in a cloud of ways-things-might-have-been… Action gives us the idea of possibility, and also an accompanying idea of dependence: *if I do this, things will go like that*… The sense of possibility thus gains an epistemic role.’ (Godfrey-Smith, 2020, 166).

Godfrey-Smith points out that a ‘sense of possibility’, a capacity of exploring a limited range of interesting alternatives to reality may have a survival value and even be found in animals:

> ‘This modal orientation linked to action might be neurobiologically deep and seen outside of humans… as rats make a spatial decision, they activate a collection of neural paths that sweep ahead of the animal’s representation of its current position, running “first down one path and then the other,” apparently representing future possibilities…’ (Godfrey-Smith, 2020, 166).

The evidence mentioned by Godfrey-Smith indicates that the capacity of selecting and realistically simulating interesting non-actual courses of future events is useful for guiding action. I suggest that one main function of imagination is to select and run such simulations. Part of the function is to develop further one course of events from a hypothesis onwards.[5] As far as there is a connection between future-directed indicatives (if… will) and past-directed subjunctives (had… would), it seems plausible that our capacity of exploring alternatives to reality extends to the latter. These are no (longer) possibilities for the future, but they are *ways things could have been*.

There is a broad range of approaches to imagination as a capacity of accurately simulating interesting alternatives to reality. For instance, imagination may play a role in mindreading, which involves simulating a situation as seen from a viewpoint that differs from one's own (Goldman, 2006). Such a simulation will be holistic. It usually includes propositional (the target subject's beliefs) and qualitative content (e.g. how the target subject feels), but its content may be exclusively propositional or qualitative (Yablo, 1993, 27; Chalmers, 2002, 151).[6] The same cognitive resources arguably are recruited in simulating objective alternatives to reality. The latter may concern what oneself or others could do, but also more detached ways things could be.

One caveat for such a view of imagination is that imagination is not *per se* constrained in the way required by realistically simulating alternatives to reality. Sceptics have questioned the use of imagination for telling what is possible and what is impossible by pointing out that we can imagine practically everything we can grasp. For instance, I may conjure up a mental image of a woman and add that she is Ruth Barcan Marcus, having just refuted Goedel's Theorem (Currie & Ravenscroft, 2002). In a sense, I thereby imagine the falsity of a truth of logic.

To counter such scepticism, I compare imagination to a tool. A tool can be used in many ways, among them dysfunctional ones, but there is also a targeted use that is more restricted. The use of imagination is sensitive to the task performed. Depending on the task, it is subject to certain implicit and explicit constraints. These constraints do not bind imagination *per se*, but conditionally. Our understanding of the task to be performed and our serious effort to perform it streamline the use of imagination. Limits may be consciously and willingly imposed, but they may also be activated in a more immediate and subconscious way.

To get a better idea of what I mean by an immediate and subconscious way, I consider Williamson's account of the folk physics backing our everyday counterfactual assessments:

> '…the folk physics needed to derive the consequents of counterfactuals such as [(ROCK) If the bush had not been there, the rock would have ended in the lake] from their antecedents may be stored in the form of some analogue mechanism, perhaps embodied in a connectionist network, which the subject cannot articulate in propositional form… the supposed premises may not be stored in a form

[5] There is a parallel to Kroedel's (2017) proposal that our reliable capacity to assess counterfactuals can be explained as enhancing evolutionary fitness. Kroedel's discussion can largely be transferred to Williamson's construal of counterfactual reasoning as an exercise of imagination, on which I am building here (Williamson, 2007, 141–165).

[6] As an anonymous reviewer reminded me, there is also a diverging understanding of imagination as essentially involving imagery or qualitative content (Kind, 2001; Kung, 2016). While I think that my use of 'imagine' is legitimate as far as there is a capacity as described, one may prefer to refer to this capacity by a different term like 'mental simulation'.

that permits the normal range of inferential interactions with other beliefs, even at an unconscious level.'(Williamson, 2007, p. 145).

When we consider whether (ROCK), we impose certain regularities concerning the trajectory of the rock. When we seriously set out to imagine what would have happened, the simulative use of imagination is strongly restrained. It may be restrained by an explicit theory, but it will often be partly restrained implicitly. One may be reliable in simulating the trajectory of the rock without being able to explicitly calculate it.

Williamson's picture is supported by results from cognitive science. According to these results, we use a quasi-Newtonian 'physics engine' in simulating counterfactual scenarios (McCoy et al., 2019, p. 237). While such a physics engine may partly be hard-wired, it must also draw on empirical information. Such information will be only partly explicit. To a large extent, it will simply be stored as dispositions how to run the simulation that are triggered by a counterfactual issue like 'what if the bush hadn't been there'.

To sum up Williamson's suggestion, the counterfactual use of imagination is restrained by the requirements of realistically simulating a scenario in which the antecedent is true. The simulation is realistic in staying as close to the actual world as permitted by the antecedent. It is honed by experience and informed by tacit knowledge. The latter is activated by using imagination in addressing a counterfactual issue. Such knowledge operates on qualitative and descriptive content and their combination. Most importantly, it is not always retrievable independently of addressing a counterfactual issue. This partly explains why counterfactuals may be useful or even indispensable in modeling.

Of course, not every part of the imaginative development has to be implicit. Especially in scientific practices that diverge from everyday reasoning, our implicit dispositions of imaginative development are subject to monitoring and correction by more regulated reasoning. This is of crucial importance when imagination is used in scientific contexts. Still, it does not follow that the immediacy described by Williamson is completely superseded by more regimented reasoning. The news value of scientific modeling as a way of reasoning from given evidence may be partially understood in terms of an interaction of implicit constraints and consciously endorsed scientific methods, beliefs and hypotheses. This has been highlighted for the comparable case of thought-experiments ever since the pioneering work of Ernst Mach (Gendler, 2007).

A general objection to the counterfactual-based approach is that counterfactual reasoning is not sufficiently well-regimented to be of much use in science. Yet when we look at alternatives like those mentioned in Sect. (1.1.) like fictions, possibilities more generally, and abstract objects, reasoning in terms of such alternatives is not guaranteed to be better regimented than counterfactual reasoning. Moreover, there is a huge literature supporting that counterfactual reasoning has a place in science. If lack of rigour is no sufficient reason to deny it a place in science, it is no sufficient reason either to deny it a place in modeling.

I have elaborated on how the use of imagination in simulating alternatives to reality can be implicitly and explicitly constrained by the task to be performed viz.

the question to be answered. I now shall elaborate on a different aspect of the use of imagination. Imagination is also involved in *selecting* models, finding relevant counterfactual suppositions to be developed in the first place, which, as Williamson emphasizes, is an ‘art’:

> ‘…selecting and interpreting models is an art – in science as well as in philosophy. It depends on good judgment, honed by experience. One must distinguish simplifications which abstract away inessential complications from those which abstract away crucial features of the phenomenon, and genuine insights from mere artefacts introduced for mathematical convenience.’ (Williamson, 2017, 169).

Imagination as described so far does more than merely allow us to assess a given counterfactual supposition.[7] It also allows us to *creatively* come up with relevant counterfactual alternatives to reality.[8]

2.1.2 The role of imagination in selecting models

With an eye on the aspect of constrained innovation, I shall now discuss the use of imagination in building and selecting models.[9]

Our sense of possibility makes us look for interesting variations of reality. Our ability to discern such variations is informed by the context of scientific inquiry. In particular, it will be informed by the scientific theories we have so far. Sometimes a model may be simply derived from a scientific theory as a perfect application of that theory. For instance, one may derive a particular model of an ideal gas from the kinetic theory of gases by adding some assumptions on the volume of the container, the number of molecules, and so on.[10] The counterfactual account applies to this case. As far as the theory makes true claims about the actual world, they figure among the candidates for being held fixed in the counterfactual model situation. Other statements of the theory can be written into the counterfactual supposition. The innovative role of imagination is less conspicuous in such an exercise of applying a theory. The framework of counterfactual imagination is still useful in settling the status of non-actual model situations: they are counterfactual situations to be developed by imagination. The imaginative development is constrained by our theory and the additional assumptions. This unified account of the status of model situations also applies when there is not yet a theory to completely guide an exercise of modeling.

[7] Williamson grants that thinking up hypotheses is a task for imagination, but his own counterfactual view of imagination comes with a tendency to downplay this task (Dohrn, 2020): ‘imagination plays a key role in science, not just in thinking up hypotheses, but in testing them.’ (Williamson, 2020). I revert the emphasis: Imagination serves not just to test hypotheses, it is crucial in thinking them up.

[8] Currie and Ravenscroft distinguish recreative and creative imagination. In the latter, the imaginer ‘…puts together ideas in a way that defies expectation or convention.’ (Currie & Ravenscroft, 2002, 9).

[9] This section has greatly profited from the comments of an anonymous referee raising issues about the role of theories, abstraction and idealization, and theoretical desirability.

[10] The kinetic theory of gases is sometimes itself described as a (more general) model. This is perfectly compatible with my account, but here I want to illustrate the aspect of derivation from a theory.

When a model is not yet completely derivable from a particular theory, the innovative role of imagination is more conspicuous. Of course, the quest for interesting variations of reality should be guided by the context of inquiry. Yet when and how to imaginatively vary reality will often be underdetermined by the explicit line of inquiry pursued. Still it is not arbitrary. It is part of a well-functioning capacity of imagination to close in on interesting variations.

As argued in the last section, our awareness of reality is accompanied by imagining various ways in which reality could be different. Some of these ways capture our attention and are developed further, depending on our aims and interests. Imagination in its primary function of exploring desirable and feasible alternatives to reality can be expected to be sensitive to several issues: what the desirable features are; whether and how they can be made real (making the antecedent supposition *A* true); what the modal status of alternative ways of making them real is; how much the different ways diverge from reality (Kment, 2014).

When we set out to imagine a model situation, our interest is special. We aim not at manipulating reality but at cognizing it. We configure the model in light of this interest. The differences notwithstanding, there remains a close connection between the two creative uses of imagination. These uses consist in developing on the one hand practically desirable alternatives to our reality, on the other hand counterfactual alternatives to our reality that are interesting for theoretical reasons. In each case, our use of imagination involves identifying interesting alternative scenarios and relating them to reality: assessing how far-fetched such alternatives are, what the relevant differences and commonalities are, and which alternatives are closer than others. Such assessments are partially implicit.

Models are not supposed to directly represent the actual world. Their function of *indirect* representation often requires a substantial departure from the actual world. The transition from reality towards a model can at least partially be characterized as a process of abstraction and idealization.[11] We focus our attention by deliberately omitting certain features of reality. We replace other features of reality by features that better suit our theorizing. The characterization can be inscribed into the procedure of imaginatively varying reality. I neither claim that all processes of abstraction and idealization involve imagination, nor that all exercises of imagination involve abstraction and idealization, but only that the exercise of model-building usually does. Imagination can recruit any mental resources. In this vein, it can also recruit the capacities underlying idealization and abstraction so as to arrive at counterfactual suppositions. While there are other exercises of idealization and abstraction for instance in forming generalizations, the counterfactual-based approach tells us what the result of this particular use of abstraction and idealization is: an imagined counterfactual situation in which some aspects of reality are simply left out and others are replaced.[12]

[11] Bokulich (2011, 41) observes that the classical trajectories in Bohr's atom model cannot be properly called idealizations of the quantum structure of electrons. She calls them fictions, but it is not obvious how fictions can be integrated into her own counterfactual-based approach. As we shall see, I account for Bohr's assumptions as simplifications dealing with ignorance.

[12] An account exclusively based on abstraction and idealization may define the result as an abstract object. I have mentioned such an account as an alternative to my own, which I cannot fully discuss. One of the

There are many ways in which looking at alternatives to our reality might support our theorizing about real target phenomena. Often the varied scenario is interesting because it combines two things: it is realistic as far as it largely preserves our theory and evidence. Yet due to some suitable variation of reality it is easier to theorize about it than about the actual world. The exemplary models from Sect. (1.2.) serve as an illustration. I can only give a highly simplified picture.

When Bohr designed his atom model, there was only very limited evidence on the precise structure of the electron shell of atoms. The evidence supported that there was a heavy positively charged atomic nucleus surrounded by a shell of light negatively charged electrons bound by electromagnetic forces, the whole system emitting light with a limited range of wavelengths. The latter correspond to the Balmer formula that had been constructed ad hoc by trial and error from the empirical data (Bokulich, 2011, 41–42).

Starting from such limited evidence and aware of his ignorance about crucial details, Bohr conjured up the following variation on given theory and data: electrons circle the atomic nucleus without losing energy on a limited number of classical trajectories, corresponding to fixed levels of potential energy. Once an electron jumps from an orbit with a higher level of potential energy to one with a lower level, a photon with an energy level corresponding to the difference is emitted.

By abstracting away from matters not yet known and stipulating others, Bohr created a scenario that was theoretically convenient as far as it conformed to the evidence available, was astoundingly simple and gave rise to powerful predictions. I shall later discuss how these predictions about a counterfactual scenario could be relevant for real target phenomena (2.3.). Bohr may not have strictly *known* the scenario to be contrary-to-fact, but it was highly unlikely to perfectly correspond to reality. It could be aptly expressed by counterfactuals.[13]

Framing the scenario as a mere model (which I reconstruct by counterfactuals) permitted Bohr to set aside the uncertainties imposed by ignorance about the precise make-up of electrons and their trajectory. It was just one of the counterfactual alternatives to the real world that conformed to the evidence. As a matter of fact it turned out to be nomically impossible yet predictively and illustratively successful.

Coming to other models, planets as point masses, water as a continuous, incompressible medium, and ideal massless-string pendula can be discussed together. In all these cases certain facts we know about the target phenomena make theorizing highly complicated: planets have unevenly distributed masses; water is not perfectly continuous and somewhat compressible; pendula have strings with masses. Mathematical theories that take into account these facts are hugely more complicated than mathematical theories for a counterfactual scenario in which they are replaced by idealizing counterfactual assumptions. Due to the mathematical idealizations, the scenarios considered are inevitably contrary-to-fact. Still they partly but relevantly match the evidence. The balance struck between matching the evidence and benefits

advantages of a counterfactual-based approach is that it facilitates the transition to reality.

[13] Counterfactuals do not require that the supposed scenario is not actual (Williamson, 2007, 137–141), yet unlike indicative conditionals, they avoid the supposition that it is a salient epistemic possibility (Stalnaker, 1975, 145–146).

in terms of simplification make them interesting for theorizing about real phenomena. Again it is to be seen how to draw lessons for the latter.

The case of the Lotka-Volterra model of population dynamics is very similar but invites some comments.[14] When one looks at graphic representations of the actual population dynamics of predators and prey, one discerns a pattern.[15] The pattern is muddled by irregularities, but it can be manifested in its pure form by idealizing, imaginatively smoothening the curve. Smoothening involves connecting discrete points representing whole numbers by a continuous curve. The smoothened curve can be obtained by letting the numbers of individuals be continuous. It is aesthetically and theoretically attractive: in contrast to the graph representing the actual development, it can be easily represented by elegant mathematical equations. The example illustrates the holistic nature of our capacity of counterfactual variation. The latter is guided by inspecting the graph and doing its mathematics. Moreover, it shows how we close in on particularly salient variations of reality. Again salience may have very different sources, sensory and cognitive ones.

There is a striking difference to Bohr's model. There seems to be something in reality itself that privileges the Lotka-Volterra model independently of our epistemic needs. Jenkins and Nolan (2012) have suggested that there is a *disposition* of populations to develop according to Lotka-Volterra equations. Populations would perfectly conform to the equations but for certain interfering factors. I shall come back to this distinction between mere epistemic convenience and a structural feature of reality itself in Sect. (2.3.).

2.2 Elaborating on a model

Once we have settled for an interesting departure from reality, the task becomes to elaborate on a counterfactual supposition that represents this departure. The supposition can be formulated as the antecedent of a counterfactual, corresponding to a question: what if things had been thus and so? Suppositions do not have to be fully propositional. We may start from some image or prop and reason 'if things were like that…'. Some things may already be logically entailed by the supposition, others have to be settled by a process of imaginatively developing the latter. My main innovation in Sect. (2.2.1.-2.2.2.) concerns the division of labour between the antecedent supposition and the consequences to be drawn.

2.2.1 Standard Counterfactual Development

I do not claim that the content of a particular model neatly divides up into the antecedent and consequent of some particular counterfactual, or that our actual reasoning

[14] I do not claim to capture the actual history of models like the Lotka-Volterra model, but only an idealized version of how modeling can proceed. Weisberg emphasizes that Volterra construed a model before looking at the more specific data (2007, 222). Yet Volterra had *some* evidence about the target phenomena to start with, in particular the observation that predators had increased and their prey had decreased due to wartime interruptions of fishing activities.

[15] For an exemplary graph plotting experimental data on different species of mites in the role of predators and prey see Huffaker (1958, 370).

process comes with such a division. I shall not try either to specify such a division in general or for particular models. Perhaps the development of a model cannot be packed into one counterfactual but will take a whole set of different counterfactuals. For simplicity, I shall consider only the case in which a single counterfactual suffices to capture a model.

Having forged a connection between the imaginative development of a supposition and the everyday language form of a counterfactual, I leave it to the semantics of counterfactuals to settle the rules of that development. I shall remain neutral on the semantics. For purposes of illustration, I refer to the simplest version of the standard minimal divergence (maximal closeness) approach (Stalnaker, 1968; Lewis, 1973):

> *a counterfactual is true precisely if the consequent C is true in all accessible antecedent A worlds that otherwise minimally diverge from the actual one.*

One advantage of the standard account is that it allows me to integrate the view that the transition between models and reality is guided by relevant similarities (Hesse, 1963; Weisberg, 2007).[16]

One worry about the possible worlds-approach is that it makes the transition from a model to reality more difficult because there is no suitable connection between merely possible worlds and the actual one. I offer two answers to this problem. The first is emphasized by Williamson: the possible worlds-approach abides by logical principles like *modus ponens*, which can be used in the transition to reality. The second answer is emphasized by Stalnaker (1968, 112): it is the *closeness* to the actual world imposed on the relevant possible antecedent worlds that makes counterfactuals informative about the actual world. If the relevant antecedent world is to remain as close to actuality as compatible with making the antecedent true, it harbours information on the actual world. That information may be used in the transition from the model to reality.

Another motive for concern about the standard account is its commitment to the vacuity of counterpossibles. If we interpret truth in a model as truth in a minimally diverging possible world, we seem to lose the advantage of dealing with impossible model situations, which I have emphasized in Sect. (1.1.). I mention three alternative ways for meeting this challenge: first, the standard account may have to be abandoned. Second, the approach may be mended by extending the range of accessible so as to include impossible worlds (McLoone, 2021). A third alternative is to accept the vacuity of counterpossibles and to resort to heuristically reasoning with *intuitions* on the truth and falsity of counterpossibles instead of their vacuous truth (Jenkins, 2010, 258; Dohrn forthcoming).

One advantage of a counterfactual-based approach is that it facilitates spelling out the connection between model assumptions and consequences. The connection of *counterfactual entailment* is somewhat stronger than logical entailment. Anything that is logically entailed by either the antecedent or the consequent of a true counterfactual would also be true if the antecedent were true. Yet many counterfactuals

[16] This includes highly specific formal proposals on how to spell out similarity as in Weisberg (2012), based on Tversky (1977).

are true without there being a logical entailment between antecedent and consequent. For instance, it may be true that (ROCK) the rock would have landed in the lake if the bush had not been there. Yet the consequent is not logically entailed by the antecedent; it follows by imaginatively developing relevant ways for the antecedent to be true (e.g. those that minimally diverge from actuality). Many truths in a model may follow in this way from the counterfactual development of relevant ways for a supposition to be made true. Moreover, the information value of a model may at least partly be explained by what we find in the course of our development instead of simply packing it into the explicit assumptions.

One criticism of the counterfactual-based approach is that the truths in the closest antecedent worlds go beyond truths in a model (Frigg & Salis, 2020). In response, truth in a model may also be understood in terms of the *relevant* counterfactuals that seem worth considering (Williamson, 2020). Moreover, truths that go beyond the limited content which we have in mind when considering a model may be relevant to the transition from the model back to reality. The overall match in fact between the closest antecedent worlds and the actual world makes a counterfactual informative about the latter.

One challenge for a closeness-based view of counterfactuals like the Stalnaker-Lewis standard approach is to deal with our practice of abstracting away from certain facts. Their negation is not part of the explicit supposition, but they shouldn't either be relevant to developing the model situation. Still they might re-enter via closeness to the actual world. One exemplary way of responding to the challenge goes as follows: in the standard approach, the weight of particular differences and commonalities is determined by a contextual similarity ordering. The similarity ordering can be settled so as to make sure that certain facts which are not just ignored but abstracted away from do not count towards closeness of model situations to actuality.

2.2.2 Minimax: sparse antecedents, rich consequents

I shall now discuss how to divide the labour between the antecedent and the consequent. Other things being equal, a theory seems the more informative the more interesting conclusions it obtains relative to the assumptions it has to make. The corresponding relationship in my counterfactual approach to models is that of counterfactually entailed conclusions and suppositions. While the suppositions are stated in the antecedent of a counterfactual, the natural place for the conclusions is the consequent. Thus, we get the maxim of doing with as sparse antecedent suppositions as possible while maximizing the interesting consequences to be counterfactually entailed by them.

The maxim just stated can be fruitfully applied to adjudicate an implicit disagreement between two main approaches to counterpossibles (counterfactuals with impossible antecedents) in modeling, those of Tan (2019) and McLoone (2021). Tan provides examples of counterpossibles to be used in modeling like the following:

> If the planets were point masses, their orbits would be….
> If there were an ideal, massless-string pendulum, it would behave….

If water were an ideal, incompressible medium, it would behave… rather than….

Here… is to be replaced by some informative description.

In contrast, McLoone considers a counterpossible on population dynamics:

(RABBIT$_{\text{McLoone}}$) If a rabbit population satisfied the logistic equation, its size would eventually equal the carrying capacity.

The difference becomes manifest upon closer inspection. Tan writes only the idealizing assumptions into the antecedent without requiring that the consequent logically follows from the antecedent. In contrast, McLoone requires that the consequent be logically/mathematically entailed by the antecedent (rather than being only counterfactually entailed). The alternative to McLoone's counterpossibles that is suggested by Tan's examples is:

(RABBIT$_{\text{Tan}}$) If rabbits were continuous in number, their population size would satisfy the logistic equation and therefore eventually equal the carrying capacity.

The reason for McLoone's stronger requirements lies in his extended closeness semantics for counterpossibles: applied to (RABBIT$_{\text{McLoone}}$), an antecedent-cum-consequent world departs less from actuality than an antecedent world in which the consequent is false; both break the laws of metaphysics, but the latter also breaks the laws of mathematics.

There are reasons for doubting McLoone's approach. He restricts the role of counterpossibles in modeling to spelling out the logical consequences of the explicit assumptions made in devising the model. We do not need counterfactual reasoning to draw these consequences. They follow by the laws of logics from the assumptions as premises. As a consequence, there is not much work for counterpossibles to do. The only remaining purpose I can discern is that they may tell us what the model description is true of: it is true of a counterfactual scenario.

The life role of the imaginative development of counterfactual suppositions gives us reason to think that counterpossibles and counterfactuals more generally make a more significant contribution. They provide guidance towards consequences that are not simply logically entailed by the antecedent. Consider (ROCK): what if the bush had not been there? Though the supposition does not logically entail that the rock lands in the lake, the counterfactual informs us about this consequence. In a similar vein, the counterfactual development of model assumptions may inform us about consequences which are not simply logically entailed. The development is the more informative the less we have to write into the antecedent and the more we can write into the consequent of a true counterfactual.

Of course, it is not trivial that idealizing assumptions already give us all the intended consequences that we take to be true in a model situation. I have proposed that we write into the antecedent just what is necessary to get the intended content of the model by *counterfactual* entailment. For instance, counterfactual suppositions like rabbit populations being continuous may not yet give us the logistic equation.

Then we have to write further assumptions into the antecedent. It does not follow that the antecedent has to logically entail what is true in the model (unless we accept McLoone's semantics).

One advantage of my maxim is the following: as I have indicated in Sect. (2.1.2.) and shall elaborate further in Sect. (2.3.), one interpretation of the Lotka-Volterra model is that actual populations have a tendency or disposition to conform to the Lotka-Volterra equations, which is superseded by interfering conditions. The same may go for a development according to the logistic equation.[17] This tendency does not become manifest when we simply write the pertinent equation as an assumption into the antecedent. Yet it transpires when we get the correct equation by counterfactual entailment from more sparse model assumptions. That entailment shows that real populations would behave according to the equation but for the failure of the idealizing assumptions to hold.

2.3 From models to reality

In developing scientific models, we aim at accounting (representing, explaining, predicting, understanding…) for real target phenomena. Yet the assumptions of models as in Sect. (1.2.) radically depart from the actual world. I interpret them as counterfactual suppositions. The decisive question becomes how they can be used for approaching real target phenomena. My innovation consists in outlining two principled alternatives. I do not pretend to cover all the varied pathways from counterfactual model situations to reality, though.

Since the motives that make us adopt model assumptions vary, I envision a corresponding plurality of pathways from models to reality. I have drawn a distinction among motives that guide us in departing from reality. Some motives have to do with our limited resources and the need to avoid overdemanding tasks. They may be called (merely) *epistemic*. Other motives may have to do with some interesting feature of the target phenomena themselves. They may be called *objective*. It is tempting to think that epistemic and objective motives for departing from reality elicit different processes of going back to reality from a model.

2.3.1 Real tendencies

I shall begin with objective motives, having to do with some feature we discern in reality. Certain counterfactuals of the form *if A were/had been the case, C would/would have been the case* used in modeling can be interpreted by discerning an underlying tendency or regularity *C* in the actual world, which becomes manifest under circumstances *A*. These circumstances *A* are contrary-to-fact. Actually, certain disturbing circumstances obtain instead. These circumstances preclude the tendency or regularity from purely manifesting itself. In this vein, one may think of model assumptions as smoothening and polishing away the disturbing circumstances.

The Lotka-Volterra model may serve as an example. Graphs showing the actual population dynamics gesture towards a regularity, but to bring out and formulate the

[17] On the relationship between Lotka-Volterra equations and the logistic equation see Blanco (1992).

regularity, we need to smoothen them by assuming populations to be continuous in number. In this case, one may hypothesize that populations have a disposition (or tendency) to develop in accordance with Lotka-Volterra equations. We may use a counterfactual to formulate the disposition. Populations would develop in accordance with Lotka-Volterra equations if only they were continuous. Interpreted as tracking a disposition, this counterfactual tells us something about reality. Suppose the minimal divergence approach to counterfactuals is correct. Then the counterfactual tells us something about reality by highlighting the situation minimally diverging from the actual one in which the lawful disposition is purely manifested as the disturbing factors do not interfere. We conclude that there is a corresponding regularity guiding the development of real populations as far as there are no disturbing circumstances.[18]

One challenge for this diagnosis is that the counterfactual situation is impossible. Jenkins and Nolan (2012) have reacted to this challenge by introducing impossible dispositions, using the Lotka-Volterra model as an example. Real populations are disposed to develop according to Lotka-Volterra equations given the impossible manifestation condition of their being continuous in number.

I shall not commit myself to impossible dispositions, just as Jenkins and Nolan avoid committing themselves to a particular view of corresponding counterpossibles. Still there is a close connection between dispositions and counterfactual conditionals. In this vein, I suggest that we take the counterfactual development of some models to tell us something about a deep feature of reality, for instance a lawful disposition of populations to develop in a certain way, and the conditions preventing or distorting its pure manifestation in reality.

To take one salient option, model counterfactuals tell us about a law. The law holds both actually and in the closest model situation, but actually there are disturbing conditions. I use ceteris paribus (cp.) conditions for illustrative purposes. Assume the law is: *cp., predator-prey populations develop according to Lotka-Volterra equations*. Now actually the cp. condition is not satisfied. Yet in the closest model situation, it is satisfied. This is ensured by the model assumptions, read as antecedents of counterfactuals. In this situation, we can discern the law: predator-prey populations develop according to the Lotka-Volterra equations. The law can be used for explanation and prediction, heeding the cp. clause. When we develop a model in order to bring out such a significant feature of reality, our motives can be called objective in the sense of having a foundation in reality.

2.3.2 Epistemic needs

From such an objective motivation I distinguish a merely epistemic one, arising from our limited cognitive and informational resources. This epistemic motivation may be illustrated by Bohr's atom model. Bohr developed the model under conditions of ignorance about the electron shell. Given the fundamental difference between classical trajectories and clouds of probability densities, we cannot say that the latter have a disposition or tendency towards the former, or that the classical trajectories

[18] It would seem weird to address the fact that e.g. rabbits are not continuous in number as a disturbing circumstance. Yet the discontinuous distribution of population sizes may be so addressed.

are the result of polishing away the disturbing noise that obfuscates some underlying law. The classical trajectories are stipulated in order to serve the epistemic needs of theoreticians. Bohr aimed at a set of simple and suggestive assumptions that allowed him to systematically derive results like the Balmer formula. Of course, the question becomes what the epistemic value of such a derivation with regard to a merely counterfactual situation is for cognizing reality. I shall come to that question in a moment.

It seems surprising that the development of a counterfactual supposition should serve so diverse needs as manifested by subjective viz. epistemic, and objective motivations of model-building. Often interpreted as 'ontic' as contrasted to 'epistemic' indicative conditionals, counterfactuals seem to better square with the objective needs of manifesting some significant distinction in reality like that between an underlying law and disturbing circumstances. My response to this challenge is to build a certain flexibility into the transition from models to reality. That flexibility should allow us to take care of epistemic motives of modeling.

2.3.3 A flexible transitional scheme

I shall draw on a proposal by Godfrey-Smith for making the transition from imagined counterfactual situations to reality. As noted before, counterfactual suppositions often cannot simply be discharged by assuming that the counterfactual antecedent is actually true. To make discharging more flexible, Godfrey-Smith develops the following scheme of reasoning:

'1. *If A then C* (…counterfactual, determined by modeling).
2. *If approximately A, then approximately C* (also a subjunctive, inferred invalidly but perhaps reasonably from (1)).
3. *Either approximately A does not hold, or approximately C* (a material conditional…).
4. *Approximately A* (via other information).
5. *Approximately C'* (Godfrey-Smith, 2020, 169–170).

Godfrey-Smith sketches a series of counterfactual variations, which guide us from a model to back to reality. As for the counterfactual (1), the antecedent *A* fixes the assumptions of the model that allow us to draw any interesting consequences for the model situation in the consequent *C*. From (1), we proceed via further conditionals, including a material conditional (3) and the 'discharging' premise that things are approximately as in the antecedent *A* of (1) to the claim that things are approximately as in the consequent *C*.[19]

While Godfrey-Smith's proposal is the most advanced and promising approach for proceeding from model counterfactuals to reality I know of, there are still several concerns to be answered. The first is what supports the transition from (1) to the

[19] An anonymous reviewer has expressed doubts that we need premise (3). Godfrey-Smith (p.c.) emphasized that his motive was to make the transition from problematic types of premises (counterfactuals) to unproblematic ones (material conditionals, declaratives) as explicit as possible.

premises hedged by 'approximately' (see Williamson, 2020).[20] A second concern is whether 'approximately' is a useful term for covering the relationship between a model and reality. One situation relevantly approximates another one, so the information on the one can be used *cum grano salis* as information on the other. In Godfrey-Smith's scheme, it is reality that approximates a model situation. Approximation can be given a precise mathematical sense. Yet the sense intended by Godfrey-Smith seems vaguer. While this vagueness gives us flexibility, still the question is to what extent the scheme of reality approximating a model situation generalizes. Often models will conspicuously depart from reality. Then things are not even approximately as in the model.

Godfrey-Smith's scheme therefore does not seem fitting when we reason with a model although we reject premise (4) and (5) of his paradigmatic reasoning process. We may use a model even if we do not take reality to approximate the model assumptions and the consequences drawn from them. Consider the following type of situation: we take the model to reveal an underlying feature of reality. The latter is so thoroughly superseded by disturbing circumstances that the model is not even approximately realized in the actual world. Although the actual world is hugely different from the model situation, the latter is informative. Again the Lotka-Volterra model may provide an example.[21] A model counterfactual like *populations would develop in accordance with Lotka-Volterra equations if only they were continuous etc.* seems to formulate a lesson on reality, just as a counterfactual *if sugar were placed in water, it would dissolve* can express a real disposition. We do not need to reason from the counterfactual to reality, the counterfactual itself tells us something about reality. Taking into account such alternatives, I have some doubts that Godfrey-Smith's scheme is generally applicable.

Still Godfrey-Smith's scheme is of partial significance. It applies when the assumptions about reality *relevantly* approximating the model situation are satisfied. This can be expected in many of those models which are chosen mainly for *epistemic* motives of theoretical convenience. We deliberately neglect divergences between reality and the model situation and thus take the one to go proxy for the other. For instance, in the case of Bohr's atom model, the probability densities of the actual trajectories of electrons are approximately like classical trajectories as far as our (Bohr's) theoretical purposes go. Other purposes the model does not serve. Applying Godfrey-Smith's template, the transition will go as follows:

1'. If the assumptions of Bohr's model held (electrons have classical trajectories), then the spectrum of light emitted by hydrogen atoms would correspond to the Balmer formula.

---

[20] A key feature of the Lewis-Stalnaker account is the resolution of the vagueness of counterfactuals by a pragmatic reasoning process of determining the contextual similarity ordering. In this vein, one may consider integrating 'approximately' into such a pragmatic process of resolving vagueness, changing the similarity ordering from one counterfactual to the next. The resulting scheme would seem more elegant while still being 'invalid but perhaps reasonable'.

[21] This is not to say that Godfrey-Smith's template does not apply to the Lotka-Volterra model. Real populations may approximate continuous ones as in the model.

2’. If the assumptions of Bohr’s model approximately held (the actual clouds of electrons in a certain respect approximate classical trajectories), then the spectrum of light emitted by hydrogen atoms would approximately correspond to the Balmer formula.
3’. Either the assumptions of Bohr’s model do not approximately hold, or the spectrum of light emitted by hydrogen atoms approximately corresponds to the Balmer formula.
4’. The assumptions of Bohr’s model approximately hold (the actual clouds of electrons in a certain respect approximate classical trajectories).
5’. The spectrum of light emitted by hydrogen atoms approximately corresponds to the Balmer formula.

The interesting steps are (2’) and (4’). They depend on an adequate understanding of ‘approximately’. This understanding harbours a substantial hypothesis of Bohr’s on which the usefulness of his model depends: whatever the actual structure of the electron shell is, it would not make a difference to those particular consequences Bohr draws from his model whether that structure or the assumed structure of Bohr’s model obtained. This hypothesis captures the risk incurred by Bohr’s model. If it is a good model, the hypothesis is borne out.[22]

Taking stock, I have advocated a certain flexibility in how to reason from model counterfactuals to reality, depending on what our theoretical interests are.[23] On the one hand, there is the opportunity of using counterfactuals to discern a certain objective feature of the real target phenomena, for instance some regularity that obtains both in reality and in the counterfactual model situation but becomes purely manifested only in the latter. On the other hand, there is the opportunity of making things more convenient for our theorizing. The transition to reality then depends on the hypothesis that the results obtained under simplifying conditions approximately hold also in reality, ‘approximately’ being sensitive to our theoretical interests. I do not claim that these two alternatives exhaust the theoretical options. I leave it to future work to further extend the counterfactual-based approach.

3 Summary

I shall close with placing my contribution in the debate. Many philosophers agree that modeling should be reconstructed in terms of counterfactuals. My new contribution consists in differentiating three stages of the use of counterfactuals in model-building.

[22] Bokulich (2011) construes the transition from Bohr’s model to reality as a case of explanation. Bohr’s model explains actual facts like the Balmer series by an isomorphism between counterfactual variations of the model on the one hand and corresponding counterfactual variations of reality on the other hand. Weighing this ambitious proposal goes beyond my scope here. I just note that, in order to assess counterfactual variations of a model situation, we have to settle the modal status of the situation. The counterfactual-based approach as developed thus naturally supplements Bokulich’s account. Bokulich’s variations can be interpreted as embedded counterfactuals: if A had been the case, then if variation B had been the case instead of A, C would have been the case.

[23] The requirement of such a flexibility is also emphasized by Weisberg (2007, 218).

In the *first stage*, we proceed from reality to a model. It involves the identification of theoretically interesting departures from reality. These departures figure in the counterfactual suppositions to be imaginatively developed. My innovation consists in a detailed account of the simulative role of imagination between creativity and restraint: imagination is used in creatively exploring interesting alternatives to reality while staying as close to reality as permitted by the supposition to be imaginatively developed.

In the *second stage*, we proceed from counterfactual suppositions via imaginative development to the consequences of these suppositions. The connection can be expressed by counterfactuals. My innovation consists in a proposal on how to divide the labour between the antecedent and the consequent of a counterfactual used in modeling: the antecedent of such a counterfactual should be kept minimal so as to derive a maximum of interesting consequences as expressed by the consequent.

In the *third stage*, we go back to reality. My innovation consists in exploring two main alternatives for the transition to reality, depending on our theoretical interests. On the one hand, these interests may involve discerning some deep feature of reality, for instance a regularity reality would conform to perfectly were it not for disturbing circumstances. Counterfactuals may be used to represent this deep feature of reality. On the other hand, we may be mainly interested in simplifying our theorizing by idealizing away from certain complications. The transition to reality may be achieved by assuming that reality approximates our idealizing model in a certain respect.

In conclusion, the imaginative development of a counterfactual supposition accounts for any stage of model-building: it leads us from reality to a model; it guides us in elaborating on the model; and it leads us back again to reality.

Acknowledgements I am grateful to the organizers of the workshop on modal modeling in science, Stockholm, May 2021, and the editors of the special issue on modal modeling for an inspiring discussion and helpful comments. My thanks also go to two anonymous referees.

Funding Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement.

Declarations

Conflict of interest None.

References

Adams, E. (1993). On the rightness of certain counterfactuals. *Pacific Philosophical Quarterly*, *74*, 1–10.

Baron, S., et al. (2017). How mathematics can make a difference. *Philosophers' Imprint*, *17*, 1–17.

Blanco, J. (1992). Relationship between the logistic equation and the Lotka-Volterra models. *Ecological Modelling*, *66*, 301–303.

Bohr, N. (1913). On the constitution of atoms and molecules, Part I. *Philosophical Magazine*, *26*, 1–24.

Bokulich, A. (2011). How scientific models can explain. *Synthese*, *93*, 33–45.

Chalmers, D. (2002). Does conceivability entail possibility? In T. S. Gendler, & J. Hawthorne (Eds.), *Conceivability and possibility* (pp. 71–125). Clarendon Press.

Currie, G., & Ravenscroft, I. (2002). *Recreative minds: Imagination in philosophy and psychology*. Oxford University Press.

Dohrn, D. (2020). Counterfactuals vs. conceivability as a guide to modal knowledge. *Philosophical Studies*, *177*, 3637–3659.

Dohrn, D. The science of counterpossibles vs. the counterpossibles of science. Forthcoming in *British Journal for the Philosophy of Science*.

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

Frigg, R., & Hartmann, S. (2020). Models in science. *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2020/entries/models-science/.

Gendler, T. (2007). Philosophical thought experiment, intuitions, and cognitive equilibrium. *Midwest Studies in Philosophy*, *31*, 68–89.

Giere, R. N. (1988). *Explaining science. A Cognitive Approach*. University of Chicago Press.

Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy*, *21*, 725–740.

Godfrey-Smith, P. (2020). Models, fictions, and conditionals. In P. Godfrey-Smith, & A. Levy (Eds.), *The scientific imagination* (pp. 154–177). Oxford University Press.

Goldman, A. (2006). *Simulating minds*. Oxford University Press.

Goodman, N. (1947). The problem of counterfactual conditionals. *The Journal of Philosophy*, *44*, 113–128.

Hesse, M. (1963). *Models and Analogies in Science*. Sheed and Ward.

Huffaker, C. B. (1958). Experimental studies on predation: Dispersion factors and predator-prey oscillations. *Hilgardia*, *27*, 343–383.

Jenkins, C. S. (2010). Concepts, experience and modal knowledge. *Philosophical Perspectives*, *24*, 255–279.

Jenkins, C. S., & Nolan, D. (2012). Disposition impossible. *Noûs, 46*, 732–753.

Kind, A. (2001). Putting the image back in imagination. *Philosophy and Phenomenological Research*, *62*, 85–110.

Kment, B. (2014). *Modality and explanatory reasoning*. Oxford University Press.

Kroedel, T. (2017). Modal knowledge, evolution, and counterfactuals. In Fischer, R.W., Leon, F. (eds.), Modal epistemology after rationalism, 179–196. Dordrecht: Synthese Library.

Kung, P. (2016). You really do imagine it: Against error theories of imagination. *Noûs*, *50*, 90–120.

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

Lewis, D. (1973). *Counterfactuals*. Blackwell.

Lewis, D. (1983). Truth in fiction. In D. Lewis, *Philosophical Papers, Vol. 1* Oxford University Press, 261–280.

Lewis, D. (1986). Counterfactual dependence and time's arrow, postscripts to 'counterfactual dependence and time's arrow'. In D. Lewis (Ed.), *Philosophical Papers, volume II* (pp. 32–66). Oxford University Press.

McCoy, J. (2019). Modal prospection. In A. J. Goldman, & B. P. McLaughlin (Eds.), *Metaphysics and Cognitive Science* (pp. 235–267). Oxford University Press.

McLoone, B. (2021). Calculus and counterpossibles in science. *Synthese*, *12*, 12153–12174.

Psillos, S. (2011). Living with the Abstract: Realism and models. *Synthese*, *93*, 3–17.

Salis, F., & Frigg, R. (2020). Capturing the scientific imagination. In P. Godfrey-Smith, & A. Levy (Eds.), *The scientific imagination* (pp. 17–51). Oxford University Press.

Stalnaker, R. (1968). A theory of conditionals. *Studies in Logical Theory American Philosophical Quarterly Monograph*, *2*, 98–112.

Stalnaker, R. (1975). Indicative conditionals. *Philosophia* 5, 269–286.

Suppes, P. (1960). A comparison of the meaning and use of models in mathematics and the empirical sciences. *Synthese12*, 287–300.
Tan, P. (2019). Counterpossible non-vacuity in scientific practice. *The Journal of Philosophy*, *116*, 32–60.
Thomasson, A. (2020). If models were fictions, then what would would they be? In P. Godfrey-Smith, & A. Levy (Eds.), *The scientific imagination* (pp. 51–74). Oxford University Press.
Toon, A. (2012). *Models as make-believe*. Palgrave-Macmillan.
Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.
Van Riel, R. (2015). The content of model-based information. *Synthese*, *192*, 3839–3858.
Walton, K. (1990). *Mimesis as make-believe*. Harvard University Press.
Weisberg, M. (2007). Who is a modeler? *British Journal for the Philosophy of Science*, *58*, 207–233.
Weisberg, M. (2012). Getting serious about similarity. *Philosophy of Science*, *79*, 785–794.
Williamson, T. (2007). *The philosophy of philosophy*. Blackwell.
Williamson, T. (2017). Model-building in philosophy. In R. Blackford, & D. Broderick (Eds.), *Philosophy's future: The problem of philosophical progress* (pp. 159–172). Wiley.
Williamson, T. (2020). *Book Review—the scientific imagination: Philosophical and psychological perspectives*. Notre Dame Philosophical Reviews. https://ndpr.nd.edu/reviews/the-scientific-imagination-philosophical-and-psychological-perspectives/
Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.
Yablo, S. (1993). Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*, *53*, 1–42.