

Biology and Philosophy* symposium on *Simulation and Similarity: Using Models to Understand the World

Response to critics

Michael Weisberg

Received: 22 December 2014 / Accepted: 12 January 2015 / Published online: 25 February 2015
© Springer Science+Business Media Dordrecht 2015

Simulation and Similarity: Using Models to Understand the World is an account of modeling in contemporary science. Modeling is a form of surrogate reasoning where target systems in the natural world are studied using models, which are similar to these targets. My book develops an account of the nature of models, the practice of modeling, and the similarity relation that holds between models and their targets. I also analyze the conceptual tools that allow theorists to identify the trustworthy aspects of models. Taken as a whole, I try to account for the ways that modeling is actually practiced by theorists, while abstracting sufficiently to understand the similarities and differences among examples of concrete, mathematical, and computational modeling.

I am grateful to Wendy Parker, Jay Odenbaugh, and Bill Wimsatt for their careful and interesting reading of my book, as well as their constructive criticisms. Although I naturally disagree with some of their critiques, I have learned much from them and appreciate the chance to clarify my own thinking about these matters. I will discuss their comments sequentially.

Parker's commentary is focused primarily on my account of similarity, what I call *weighted feature matching* (WFM). This account formalizes the idea that models stand in representational relations to their targets in virtue of sharing some set of highly important features with their targets, not lacking too many of these features, and not having too many extra features. Scientific context, as recorded in what I call the modelers' construal, determines the choice and weighting of important features. Parker argues that I equivocate on the aim of giving such an account, and then offers some very interesting technical objections about the content

M. Weisberg (✉)

Department of Philosophy, University of Pennsylvania, 433 Cohen Hall, Philadelphia, PA, USA
e-mail: weisberg@phil.upenn.edu

of the feature set, the notion of sharing a feature, and the assignment of weights to features.

What is Weighted Feature Matching an account of?

What is WFM an account of? Parker suggests three possibilities: (a) an account of the model/world relation in virtue of which scientific models are successful, (b) an account of the relation that generally holds between models and their targets, or (c) an account of the judgments of scientists about the relationship between models and their targets. She argues that my discussion sometimes moves between these three aims, and that I have conflated them. While I discuss all three of these motivations, I don't think my account conflates them. Rather, I see my account as addressing all three of these themes.

I begin from Cartwright (1983), Giere (1988), Teller (2001) and Godfrey-Smith's (2006) insight that good scientific models are similar to their targets in certain respects and degrees. WFM is a way of filling in what this claim amounts to. The main components of the account are a set of features, weights on those features, and an abstract, set-theoretic expression describing the similarity relation. The contents of a feature are Cartwright, Giere, and Teller, and Godfrey-Smith's "respects," while the weighting function gives the "degrees."

This is of course all very vague, so let me elaborate on some details. In its most abstract form, with no terms filled in, the WFM expression describes the general form of similarity relations that can hold between models and targets. It says that:

$$S(m, t) = \frac{\theta f(M_a \cap T_a) + \rho f(M_m \cap T_m)}{\theta f(M_a \cap T_a) + \rho f(M_m \cap T_m) + \alpha f(M_a - T_a) + \beta f(M_m - T_m) + \gamma f(T_a - M_a) + \delta f(T_m - M_m)}$$

where m and t correspond to the model and target, M and T the sets of features possessed by the model and target that are members of the feature set Δ , f a weighting function, and the additional greek letters correspond to weights on each term.

With no parameters set and no weighting function defined, the equation describes an infinite set of potential relations—at least one of which almost certainly holds between a model and a target. So it is correct to say that WFM is an account of the sort of relation that generally holds between models and their targets, which was Parker's second option. But the story doesn't end here.

Parker notes that I sometimes adopt psychological terms in my discussion. For example, I say that my account of similarity "should also be able to help diagnose such extraempirical disagreements, locating the sources of disagreement in context, use, and weighting of various features of the model." I also say that the account

should reflect judgments that scientists can actually make, as opposed to asserting that the relation holds between inaccessible, hidden features of models and targets. This is clearly a modal desideratum, because in many

cases theorists won't necessarily articulate the grounds for the judgments of similarity—the judgments are just made. Nevertheless, when it matters, such as in cases of disagreement, theorists should be able to work out the grounds for their similarity judgment (137)

and I go on to note that this is similar to Grice's calculability assumption (Grice 1975, 1981).

Bring these psychological ideas to my account might look like a confusion, but I think it makes sense for two reasons. First, even the most strictly metaphysical parts of my account have a substantial pragmatic element. This is because on my view, the relation between a model and a target depends, in part, on the scientific context. Further, and this is the point of the passages quoted above, my account allows disagreements to be articulated and characterized. Although judgments about similarity relations are often made implicitly, in cases of disagreement among scientists, they can be articulated explicitly. This means they are in principle cognitively accessible.

What goes in a feature set?

Parker's second criticism concerns my notion of a feature set. To understand her critique, let me review some of the details of my account: WFM requires defining a weighting function over the feature set Δ , the set of features with respect to which a model is compared to a target. Given the elements of this set, the weighting function tells us how similar a model is to its target. When WFM is coupled to my overall picture of modeling, I see the choice of a feature set as equivalent to the choice of the modeler's intended scope, which is connected in important ways to the modeler's choice of target (Elliott-Graves and Weisberg 2014).

A specification of intended scope is a specification of what parts of the real-world one intends to capture, and this corresponds to some set of features of interest. Even though target systems are abstract relative to real-world systems, they will still be quite rich compared to many models since they are parts of real-world systems. So feature sets will be quite rich, even while excluding features of real-world systems.

Parker wonders about the relationship between my discussion of feature set construction, and my discussion of different types of modeling practices. For the latter, I say that a particular type of practice may be such that some of the term weights for the WFM equation are set to zero. For example, in how possibly modeling, we are interested in overlap between the attributes of model and target, but not the mechanism. Models will contain some mechanism or other by which the systems gains its attributes, but the modeler is not trying to find the actual mechanism by which this happens (if she finds it, so much the better, but that isn't the point).

If we suppress the weighting function and term weights, the WFM equation for how possibly modeling has this form:

$$\frac{|M_a \cap T_a|}{|M_a \cap T_a| + |M_a - T_a|}$$

This means that a model is similar to its target to the degree that it has some mechanism or other that can reproduce the attributes of the target. The attributes of the model and target must be similar, but any plausible mechanism can be used to generate these attributes.

Parker sees a dilemma:

Either the goal of how-possibly modeling is such that its feature set never includes mechanisms or, even when the goal is how-possibly modeling, the feature set can sometimes include mechanisms. If it never includes mechanisms, then there is no need for the weighting parameters, because M_m and T_m are always empty (\emptyset), and according to Weisberg $f(\emptyset) = 0$. If it sometimes does include mechanisms, then not all features in the feature set are important for establishing similarity, which directly contradicts Weisberg's claim that they are.

I don't accept this dilemma. As I said above, the choice of the feature set is equivalent to the choice of the target system, not the model, and certainly not the goal of modeling. Depending on how the target system is specified, the feature set may include (or not) mechanistic features.

However, the fact that the theorist is engaged in how possibly modeling *does* set constraints on the overall weighting function. The most perspicuous way to see this is that how-possibly modeling puts zero weight on the mechanism terms. But one can also see this as giving weight of zero to mechanistic features individually. So while only the specification of the target determines the content of the feature set, the modeling activity (how-possibly vs. minimal vs. hyper-accurate modeling) determines how the weights are set.

Do model and target need to share features or have similar features?

Parker's third criticism has to do with my account of feature matching. She thinks that it is odd to say that one is giving an account of similarity, which admits of degrees, on one hand, and that features must be shared on the other. So she concludes that I think sharing of individual features comes in degrees, and draws on a passage where I say something that can be interpreted along these lines. Parker concludes that

This sort of 'matching' can be cast in terms of genuinely shared properties but only, it would seem, in a rather awkward way: The Bay Model and the real Bay share the property of having a Froude number that is within 0.1 of the real Bay's number. It is more natural to say that the Bay model and the real Bay have *similar* Froude number ...

Perhaps it is more natural to express the relationship this way, but this is not my account, nor need it be. WFM requires that features be shared, but there is no reason that features can't take the form of ranges.

There are many different ways to partition up properties of models and targets. For example, if we say that model and target both have Froude number of 0.1, does

that mean 0.10000 or 0.100 or 0.100000000 or some other degree of precision? Each of these is really a range. The number 0.100 usually means 0.104–0.095. So to specify a feature as the sharing of Froude number within 0.1 is just code for saying that if the target's Froude number is exactly 1.0, then the model's value has to be somewhere between 1.1 and 0.9 to share the feature with the target.

This is *how* I can allow matching between features that are not perfectly precise. But what of the larger issue that Parker raises? Isn't it weird that an account of similarity admitting of degrees bottoms out with features that are either shared or not?

I argue no. Critics of the use of similarity relations such as Quine (1969) and Goodman (1972) ask us for precisely this type of analysis. They say that similarity is illegitimate unless we can give a reductive analysis of it. To do so isn't to change the character of similarity, but rather to explain the character of similarity.

Can feature weights be set independently in WFM?

Parker's final criticism concerns the *independence assumption* of WFM. This is a restriction that I introduced because in Tversky's original account, a weight had to be assigned to each element of the *power set* of the feature set. Since the cardinality of the powerset is 2^n for sets of cardinality n , this means that the weighting function would be astonishingly complex for even moderately sized feature sets. On the other hand, if we impose an additivity restriction roughly of the form $f(A) + f(B) = f(A, B)$, then the weighting function only needs to be defined over n terms. This restriction is motivated by my contention that for the account to be plausible, scientists need to be able to work out their weighting functions if called on to do so.

Parker raises a very interesting potential counterexample:

Suppose a modeler is interested in predicting quantity C with a specified level of accuracy, and she believes that her model will succeed in doing this if either (1) insofar as it underestimates A , it does so by an amount that is compensated for enough by an overestimate of contributing quantity B or (2) insofar as it overestimates A , it does so by an amount that is compensated for enough by an underestimate of B . It is not entirely clear how fidelity criteria related to A and B will be specified here—whether the criteria will be disjunctive, etc.

No doubt there are cases where setting feature weights independently is difficult or impossible, perhaps in the kinds of climate forecasting models that Parke (2014) has extensively discussed. However, I don't think the case described above is a counterexample. In a case like this, our primary fidelity criterion would be prediction of C with accuracy α . This would be reflected in the weight given to feature C 's being shared between the model and target. Over-estimating A or under-estimating A is only indirectly relevant to the theorists' goals and would not be given a high weight. The point is that only the features whose presence is essential are given a high weight. Means to finding them, even when known, should be given lower weights.

However, there may be cases where there are genuine tradeoffs between the achievement of different features (whose presence in the model are the goal of modeling), so that it only makes sense to consider their weighting functions together. In those cases, Parker is certainly right that good weighting functions have to take this into account. Perhaps in these cases we should see the independence assumption as relaxed, or, perhaps, we could construct compound features of the tightly coupled ones, and weigh them accordingly.

Must I be a mathematical realist (and if so, is that incompatible with WFM)?

Odenbaugh focuses on a different set of issues, specifically the ones surrounding the ontology of models. He specifically critiques my agnosticism about the ontology of mathematics, my distinction between models and model descriptions, and my handling of the debates about fictionalist accounts of models.

Odenbaugh argues that I cannot remain agnostic about matters of ontology because my insistence on the distinction between models and their descriptions commits me to mathematical realism. Only the mathematical realist, he argues, can claim that equations or other model descriptions are descriptions of something, namely mathematical models. But this option isn't very appealing, he argues, because mathematical realism is incompatible with WFM. If models are genuine mathematical objects, they can't have physical properties like a period, length, and so-forth. Since my position requires models to share features with their targets, and targets do have physical properties, it is incoherent.

I think this objection misses the mark in substantial ways. First, as I say throughout the book, I try to articulate the main claims of my book at what Stacey Friend dubbed the *epistemic level* of theorizing. This level of theorizing asks what categories and concepts we need to construct an account of the practice in question, not what the ultimate ontology has to look like in order to provide a supporting structure for this account. So when I say I want to be neutral about the ontology of mathematics, it is because I am trying to give an account that can make sense of the scientific practice as it currently stands, not make ultimate claims about the ontology of models.

Like most philosophers of science, I assume that whatever account of ontology of mathematics is ultimately shown to be correct, it will have to be compatible with all or most aspects of successful scientific practice. So whatever account of mathematical ontology is true will have to let us make sense of apparently realist talk about mathematical objects.

Even setting the deeper questions of ontology aside, another part of Odenbaugh's objection requires response. Mathematical objects as understood by scientists don't have properties that would make them similar to real-world targets, and they have many properties that no physical system can have. This is an important objection when directed at those who see mathematical models as strictly mathematical objects, such as some structural realists and traditional defenders of the semantic view of theories. But I think that mathematical models are *interpreted* mathematical objects. A harmonic oscillator model can be said to have a period because modelers

interpret part of its mathematical structure as denoting a period. These relations of denotation are such that it makes sense to say that the model, but not the mathematical structure itself, has properties like a period.

Are models constructed by their descriptions?

Odenbaugh challenges my claim that models are created by their descriptions. Having had the chance to reflect further on this claim, I agree with his critique; this is not a good way to put the point I wanted to make. Model descriptions represent models, they do not create them. Models are created when modelers construct a structure or choose a pre-existing structure, and then interpret this structure. Sometimes writing down a model description is an essential part of this process, such as in the creation of ODD model descriptions for computational models (Grimm et al. 2010). But even in these cases, there is almost always a developmental cycle, where the modeler refines both the model and the description of the model.

Is the problem of variation really a problem?

Most of Odenbaugh's other objections to *Simulation and Similarity* concern my handling of fictionalist accounts of mathematical models. Although fictionalist accounts differ in various respects, most proponents of fictionalism argue that mathematical models are not mathematical objects but are instead imaginary systems that would be concrete if they are real. In my book, I discuss various problems for the specific accounts, and then give a number of objections to the idea in general, preferring, as I do, to see mathematical models as interpreted mathematical structures.

I argue that one reason to rejection fictionalist accounts of mathematical models is because of what I call the *problem of variation*. While in literary fictions, variations in the way that people imagine a fictional world can be part of what makes fiction enjoyable, the content of scientific models shouldn't vary between users of the model. Mathematical models should be completely public in the way the appreciation of fictional worlds is not.

Odenbaugh responds by saying that there is no more reason to think that there is problematic variation for models than for literary fictions. He doesn't explain why he thinks this, but I suspect it is because in both cases, the important details are either mentioned in the explicit communication of the model/story or else we have principles of generating the unmentioned properties.

To see why I don't accept this line of response, we need to distinguish between focal and non-focal properties of fictions. Focal properties are properties of the fictional world that are required to understand the story (e.g. Rohan is west of Mordor), non-focal properties are other properties required to fill in the details of a fictional world, but that aren't necessary for understanding what is happening (e.g. the number of toes on Orc's feet). Most fictional texts leave both focal and non-focal

properties unmentioned, and descriptions of models are especially sparse in comparison to works of literary fiction.

If we understand models as concrete fictional worlds, they are sure to have many focal properties not explicitly given by their descriptions. But if that is right, then we need to account for how there can be focal properties of models that are generated but not contained in their descriptions. Odenbaugh seems confident that the same ways we deal with the unmentioned focal properties of literary fictions will be sufficient to deal with models, but I am fairly confident that they will not.

A popular way to deal with unmentioned focal properties in literary fiction is to appeal to Lewis (1978) and Walton's (1990) mutual belief principle. This principle says that we import into a fictional world all of the beliefs that would be accepted by people in that world. While this is a very plausible account of how unmentioned focal properties of fictional stories can be filled in, it is unlikely to work in the case of mathematical models. Many of the relevant properties are simply unknown to the scientific community before they are researched.

For example, consider the Lotka-Volterra model. The mathematics of the model are given at the population level, so spatial arrangement of individuals is unmentioned in the model description. On the non-fictionalist view of mathematical models, this means that the mathematical model is literally about populations, not individuals. It isn't that the model doesn't mention particular individuals; it isn't about them at all. But if the model is a concrete fictional scenario, then we have to understand populations concretely. Since real, concrete populations are composed of individuals, the Lotka-Volterra model is actually about individual organisms located somewhere in space.

Modern research using agent-based models (Weisberg and Reisman 2008) shows that when we represent populations as being composed of individuals, the Lotka-Volterra model is sensitive to spatial arrangements. Some spatial arrangements generate its characteristic properties, and some of them won't.

Does this mean that for the fictionalist, only some concrete, imagined populations are Lotka-Volterra populations? It must be, because how can concrete instantiations that don't have these properties be instantiations of the model? If that is the case, the mutual belief principle would have to rule out the instantiations that don't have these properties. Since nothing in everyday life or biology tells us this, this seems to be an ad hoc application of the principle.

Can fictionalists account for different representational capacities?

A second objection of mine to the fictionalist view is that it cannot account for the differing representational capacities of different kinds of models. I write that:

models can be discrete or probabilistic, aggregative or individualistic, spatially explicit or not, and so forth. If models are mathematical objects, these differences are easy to make sense of. Different kinds of models will use different kinds of mathematics and this will account for differences in their representational capacities. However, fictions accounts cannot make these distinctions. (61)

The reason I don't think fictions can make these distinctions is because concrete fictional scenarios, at least in biology, are one in which some determinant thing happens to individuals. Odenbaugh responds that I misunderstand the fictionalist account and that the "fictionalist makes-believes that the model description is true." But this is just what I claim is hard to understand in the kinds of examples I give. How does a probabilistic outcome get represented in a concrete fictional scenario? Are there also modal facts internal to the scenario? How are they evaluated? And how about a model of an infinite population, or an ensemble of infinite populations?

What is the face value practice of modeling?

Odenbaugh's third objection concerns the *face value practice* (Thomson-Jones 1997) of modeling. Godfrey-Smith and other fictionalists say that the face value practice suggests modelers are conceiving of models as fictions, where I argue that there is too much variety to draw that conclusion.

Odenbaugh agrees that there is variation, and that modelers don't always start their discussions with the locution "Imagine that..." But he also points out that storytellers don't always start with this phrase either. So a variety of locutions, and perhaps even cognitive styles, doesn't threaten the face value practice and its support for the fictionalist account.

I strongly disagree with Odenbaugh. There is a huge variety of structures that can serve as mathematical models. To take one important example: mathematical structures used in modeling are often high dimensional. These dimensions might represent loci, population level properties, space and time, and many other things. Whatever they represent, beyond a certain point, such models cannot be imagined in any real sense. Simple systems can be imagined, but complex structures can only be manipulated mathematically or inside of a computer.

If one thinks that we ought to give a very deflationary account of modeling, where to model is simply to represent a system in an idealized way, than this argument loses its force. But if one wants to characterize modeling as Godfrey-Smith, Frigg, and I do, as a kind of surrogate reasoning, then this is not an attractive way to set things up. Modeling involves the construction and analysis of something, so what kind of thing is it? I argue that it cannot be an imaginary, concrete system in these kinds of cases.

How much is modeling like experimentation?

The main theme of *Wimsatt's* comments concerns the parallel between modeling and experimentation. He contends that my account, especially of model/target relations and model analysis, is directly applicable to experiments. Moreover, he thinks my account shows that the differences between models and experiments seems to be "strikingly trivial" a matter of "medium" rather than "strong formal similarities."

I am sympathetic to just about everything that Wimsatt says, and admit this is a theme which I have not spent enough time thinking about before writing this response. Let me start with a small disagreement: I think that asking about the relationship between models and experiments is a category mistake. Models are objects (interpreted structures), experiments are procedures for analyzing objects. Simulations and mathematical analyses are also procedures for analyzing objects. Rather than distinguish between models and experiments, we should ask about how the procedures involved in modeling and experimentation are related.

We can thus reframe Wimsatt's question as follows: What, if anything, is the difference between modeling and experimenting? When we reframe the question this way, I think we should accept Wimsatt's perspective. Many activities involving concrete models are experimental, and many important experiments seem to involve models. As Wimsatt points out, the calibration and analysis of the Bay Model involved, quite literally, experimentally manipulating the model. Similarly, in canonical experiments, we often construct a model of the system we really want to study. Since it is difficult, for example, to study natural selection in the wild, many evolution experiment evolution experiments are done with model organisms such as fruit flies and *E. coli*. So when we ask the question this way, and think about concrete models, it is hard to see where a line can be drawn.

Things are more complex when we think about computational and mathematical models. There is no totally straightforward analogy between experimentally manipulating something in the laboratory and the manipulation of a mathematical or computational model by hand or using a computer. Nevertheless, there are striking parallels. With experimental systems, one aims to systematically manipulate the important variables in order to understand how the properties of the system depend on those variables. Similarly, theorists study models in order to understand how some of the model's properties are related to other properties. In the case of computer simulations, this can look very much like experiments: dependent variables and parameters are set, and the time course of a system is studied.

All of this is not to say that there aren't differences between the kinds of activities that scientists call modeling and the ones that they call experimenting. One difference between many modeling activities and many experimental activities might be understood using Simon's (1969) distinction between the natural and artificial. Modeling tends to involve the study of objects which are artifacts, while experimentation tends to involve the study of natural systems, or objects that have their origins in natural systems. In an appropriate hedged sense, this is a reasonable thing to say. But it is important to note, as Wimsatt does, that many experimental systems are highly artificial in the sense that while they are derived from nature, they would not be found in nature but for the manipulation of scientists (Morgan's fruit flies are an excellent example). Similarly, some models systems are natural occurrences. In my book, I discuss several examples of natural experiments which are used as concrete models. So while once can see experimenting and modeling as often on different parts of the natural/artificial continuum, this is just what it looks like most of the time.

Another difference between many instances of modeling and many instances of experimentation involves theoretical aims. Most instances of experimental work and

some instances of modeling aim to explain how a particular system works, including what mechanisms drive its behavior, how the behavior can be changed, and what the system will do in the future. Some, and classically much of, modeling effort has had a different, theoretical aim. It aimed to articulate a systematic framework to account for a broad range of phenomena, study phenomena not known to exist, learn about the formal properties of systems, characterize the relationship between the system of interest and possibly related systems, and so-forth. The hundreds of papers about Schelling's model of segregation are of this character, and it is hard to think of equivalents from the experimental literature. This difference between many of the activities called modeling and experimentation is, again, not a hard and fast distinction. But I think it is a real difference.

One last small point of disagreement with Wimsatt. At the end of his comments, he suggests that we might be able to distinguish experiments and models in material terms: experimental systems are made of the same kind of stuff as their targets. Like Parke (2014), I think this view is mistaken because experimental systems are often not made of the same kind of stuff as their targets, and sometimes models are. The Bay Model is, in part, made of salt water. The *E. coli* in experimental evolution experiments are not the same as any organisms in nature, but more importantly, are often standing in for very different kinds of organisms, and even sometimes for macroevolutionary trends (Blount et al. 2008).

Robustness as a general tool

Another theme in Wimsatt's commentary concerns robustness analysis. He notes that the notion of robustness I discuss in *S&S* is more narrow than the one he endorses. The scope of Wimsatt's own notion of robustness is as follows.

[A]ll the variants and uses of robustness have a common theme in the distinguishing of the real from the illusory; the reliable from the unreliable; the objective from the subjective; the object of focus from artifacts of perspective; and, in general, that which is regarded as ontologically and epistemologically trustworthy and valuable from that which is unreliable, ungeneralizable, worthless, and fleeting (Wimsatt 1981).

I fully accept the idea that there is a family of procedures reasonably called robustness analysis that involve finding conditions of reliability for systems, be they concrete, computational, or mathematical, and whether they are natural or artificially occurring. This broader notion involves gaining confidence in a robust result.

We can call all of these things robustness, but I also think that the more narrow version of robustness that I discuss poses special challenges. Unlike finding the conditions under which a circuit will fail or a bridge will collapse, finding a robust theorem is a matter of gaining understanding of a real or potentially real-world system using models known to be inaccurate. This is why Levins characterizes robust theorems as the truth "at the intersection of independent lies" (Levins 1966). Many of the procedures (varying parameters, structure, representational systems)

are common between these activities, but the goal is different in the two cases. Everyone agrees that gaining confidence in a system in the region of certain parameters is something that we can do, but it remains controversial that we can understand something about the world through robustness analysis. As Wimsatt himself once wondered (1987), how can false models be a means to truer theories? Robustness analysis is a big part of the answer.

Acknowledgments I am grateful to Shereen Chang, Devin Curry, Louise Daoust, Alkistis Elliott-Graves, Karen Kovaka, Dave McCandlish, Emily Parke, and Carlos Santana for a very stimulating discussion about the symposium papers and about what I might say in response.

References

- Blount ZD, Borland CZ, Lenski RE (2008) Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci* 105(23):7899–7906
- Cartwright N (1983) *How the laws of physics lie*. Oxford University Press, Oxford
- Elliott-Graves A, Weisberg M (2014) Idealization. *Philos Compass* 9(3):176–185
- Giere RN (1988) *Explaining science: a cognitive approach*. University of Chicago Press, Chicago
- Godfrey-Smith P (2006) The strategy of model based science. *Biol Philos* 21:725–740
- Goodman N (1972) Seven strictures on similarity. In: *Problems and projects*. Bobbs-Merrill, Indianapolis
- Grimm V, Berger U, DeAngelis DL, Polhill JG, Giske J, Railsback SF (2010) The ODD protocol: a review and first update. *Ecol Model* 221(23):2760–2768
- Grice HP (1975) Logic and conversation. In: Cole P, Morgan J (eds) *Syntax and semantics: speech acts*, vol 3. Academic Press, New York, pp 41–58
- Grice HP (1981) Presupposition and conversational implicature. In: Cole P (ed) *Radical pragmatics*. Academic Press, New York, pp 183–198
- Levins R (1966) The strategy of model building in population biology. *Am Sci* 54(1):421–431
- Lewis D (1978) Truth in fiction. *Am Philos Q* 15:37–46
- Parke E (2014) Experiments, simulations, and epistemic privilege. *Philos Sci* 81(4):516–536
- Quine WV (1969) *Natural kinds*. In: *Ontological relativity, and other essays*. Columbia University Press, New York, pp 114–138
- Simon HA (1969) *The sciences of the artificial* (136). MIT Press, Cambridge
- Teller P (2001) Twilight of the perfect model. *Erkenntnis* 55:393–415
- Thomson-Jones M (1997) Models and the semantic view. *Philos Sci* 73:524–535
- Walton KL (1990) *Mimesis as make-believe: on the foundations of the representational arts*. Harvard University Press, Cambridge, Mass
- Weisberg M, Reisman K (2008) The robust volterra principle. *Philos Sci* 75:106–131
- Wimsatt WC (1981) Robustness, reliability, and overdetermination. In: Brewer M, Collins B (eds) *Scientific inquiry and the social sciences*. Jossey-Bass, San Francisco, pp 124–163
- Wimsatt WC (1987) False models as means to truer theories. In: Nitecki M, Hoffmann A (eds) *Neutral models in biology*. Oxford University Press, Oxford, pp 23–55