
Article:
Stuart, Mike orcid.org/0000-0002-4165-2641 and Wilkenfeld, Daniel (2022) Understanding metaphorical understanding (literally). European Journal for Philosophy of Science. 49. ISSN 1879-4912

https://doi.org/10.1007/s13194-022-00479-5

European Journal for Philosophy of Science (2022) 12: 49
https://doi.org/10.1007/s13194-022-00479-5

PAPER IN GENERAL PHILOSOPHY OF SCIENCE

Understanding metaphorical understanding (literally)

Michael T. Stuart [1,2] · Daniel Wilkenfeld [3]

Received: 5 April 2021 / Accepted: 11 May 2022 / Published online: 19 July 2022

Abstract

Metaphors are found all throughout science: in published papers, working hypotheses, policy documents, lecture slides, grant proposals, and press releases. They serve different functions, but perhaps most striking is the way they enable understanding, of a theory, phenomenon, or idea. In this paper, we leverage recent advances on the nature of metaphor and the nature of understanding to explore how they accomplish this feat. We attempt to shift the focus away from the epistemic value of the content of metaphors, to the epistemic value of the metaphor's consequences. Many famous scientific metaphors are epistemically good, not primarily because of what they say about the world, but because of how they cause us to think. Specifically, metaphors increase understanding either by improving our sets of representations (by making them more minimal or more accurate), or by making it easier for us to encode and process data about complex subjects by changing how we are disposed to conceptualize those subjects. This view hints towards new positions concerning testimonial understanding, factivity, abilities, discovery via metaphor, and the relation between metaphors and models.

Keywords Metaphor · Understanding · Factivity · Imagination · Models · Creativity

✉ Michael T. Stuart
mike.stuart.post@gmail.com

[1] Institute of Philosophy of Mind and Cognition, National Yang Ming Chiao Tung University, Taipei, Taiwan

[2] Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, London, UK

[3] School of Nursing, Department of Acute and Tertiary Care, Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA, USA

1 Introduction

Metaphors play important roles in all stages of scientific practice. They can be used to justify the pursuit of one theory over another (Herrington & Jablonka, 2020; Nyrup, 2018); they can help scientists see things in new ways and thereby come to make new discoveries (Jacob, 2001; Nersessian, 1984; Spranzi, 2004); they can justify theoretical claims and provide explanations (Hills, 2016; Levy, 2020); facilitate conceptual change and commensurability (Nersessian, 2015); and mediate the interpretation of hard-to-grasp theoretical structures (Stuart, 2016, 2018). It is perhaps due to these affordances that metaphors are omnipresent in both science education (Aubusson et al., 2006) and science communication (Beger & Jäkel, 2015; Kendall-Taylor et al., 2013).

One way to capture several of the above contributions is in terms of *understanding*. Scientists, as well as science consumers, frequently come to understand complex phenomena by way of metaphorical appeals to elements of the world with which we are more familiar. The market is governed by an invisible hand, genes propagate because they're selfish agents, the brain encodes and decodes information, etc. These metaphors all have their detractors, but few would deny that they were helpful at some point for understanding a target system.

Some may find it puzzling that metaphors can be used to increase understanding. Arnon Levy (2020) traces this puzzlement to metaphors' literal falsity, historical propensity to mislead, and the fact that they seem to exist in a fundamentally unconstrained and lawless place. Nevertheless, metaphors have played and continue to play a crucial role in scientific development—given their aforementioned features, it is worthwhile to ask how they perform this trick.

This paper builds on recent work by Elisabeth Camp and Arnon Levy, who connect metaphors to scientific understanding via imagination. In the next section we explain some interesting features of these accounts, and then point out a tension: either metaphors are asserted indicative statements that carry information about their targets, or they are *invitations* to look through a certain lens. We resolve this tension by allowing metaphors a primary and secondary function, and allowing speaker intention to determine which function is primary. Whichever function is primary, what seems to matter most for an analysis of metaphorical understanding are the consequences of the metaphor. We build on this idea using an account of understanding to define the states of understanding promoted by a consequentialist epistemology of understanding.

2 Features of metaphors

In this section we want to summarize a number of key insights made about metaphors, all of which we find plausible. A common starting point is the work of Kendal Walton, who characterizes metaphors as remarks that suggest, imply, introduce, or call to mind games of make believe. In those games, we are invited to make believe as though the sentence uttered were true (Walton, 1993). We can point out a number of complications here without casting doubt on the basic idea that metaphors are invitations to make believe. For one, some metaphors are literally true (e.g., "no one is an island"). It's not clear why we would need to play a game of make believe in which the sentence uttered is literally true, since it *is* literally true. Second, it does not seem that scientists are keen

to have people make believe their metaphorical claims are literally true. For example, Dawkins does not invite us to pretend that genes are *literally* selfish agents (he is explicit about this). To make believe that genes are literally selfish agents, or that markets are literally guided by invisible hands, would lead us too far astray from the real features of the target system in which we are interested.

Still, we can retain the idea that metaphors are invitations to do something. According to Max Black, metaphors invite us to see one thing through the lens of another, e.g., by suppressing some details and emphasizing others (Black, 1962, 41).[1] Building on this insight, Camp argues that metaphors invite us to assume a certain taxonomy (telling us what entities to focus on); they direct our attention toward certain features of an object (often through vivid imagery); and they tell us which features ought to matter, and how much they matter (Camp, 2020, 309–10). According to Camp, how much a feature matters will depend on how *prominent* it is in a classification scheme, which is a measure of taxonomy-relative *diagnosticity* and *intensity*. Features also matter more or less depending on how *central* they are. Centrality is a function of a feature's connectedness to other features, which can be measured counterfactually: the more things change upon removal of a feature, the more central it is. Those connections can be conceptual, emotional, aesthetic, or explanatory.

In sum, for Camp, metaphors effect changes in our perception, thought, and action by causing us to bring a new *perspective* to bear on an object (Camp, 2006, 2009, 2020). When thinking metaphorically, we take one perspective normally applied to *X*s, and apply it to *Y*s. Perspectives are sets of dispositions to treat objects in certain ways. For example, framing genes as *selfish* causes us to apply the perspective we have about selfish agents to genes. In this case, our selfish agent-perspective is the set of dispositions we have when it comes to selfish agents, which causes us to predict that selfish *genes* (like other selfish things) will hoard available goods for themselves, manipulate competitors to get what they want, and so on. Importantly, the dispositions that make up a perspective are a necessarily open-ended set: tomorrow we might experience something new which alters that set.

3 Understanding and the nature of metaphors

Philosophers disagree about what kinds of understanding there are, and which are the most interesting or fundamental. Views that associate understanding with grasping true explanations are seen as somewhat more traditional, while those that emphasize

[1] One popular way to think about metaphors as lenses is by reference to the imagination. Levy writes, "metaphors engage the imagination. They are a type of figurative device, imposing an imaginative description on a real-world target" (Levy, 2020, 292). For Camp, the kind of imagination relevant is a Kantian synthetic sort, which unites disparate elements into a coherent whole in experience (Camp, 2020). Importantly, as Camp points out, there are different "directions" of imagination that might be relevant here. For example, we might imagine *of* a real market that it was guided by an invisible hand, and compare that to what we know about that real market's behavior. Or, we might imagine a fictional version of a real market that has always been (in the fiction) guided by an invisible hand, and compare that to the real market's behavior. It could be objected that we do not always *feel* like we are imagining when we process a metaphor, but the strength of this objection depends on the sort of imagination we have in mind, e.g., if subconscious imagination is possible, then this worry loses much of its bite since we can imagine without realizing we are doing so (see Stuart, 2019, though see Kind, 2021).

abilities are seen as more radical. We'll begin with the more traditional viewpoint, on which we increase understanding by representing a target in a way that satisfies at least two criteria: 1) the representation has some good-making features, such as relative accuracy, coherence, or being "true enough" (Elgin, 2017); and 2) the representation should be connected in the right way to the understander. That is, the agent must *grasp*, *possess*, *accept,* or *believe* the content of the representation, as well as the relation between the representation and the explanandum. The representation might take the form of (part of) an explanation, an argument, an argument pattern, a system of equations, a model (Dellsén, 2020), or an "explanatory nexus" (Khalifa, 2017).

More generally, structures conveying understanding (whether they are metaphors, models, explanations, diagrams or arguments) *represent* a target in an epistemically useful way by carrying information about things like laws of nature, dispositions, causes, mechanisms, dependency relations, or probabilities, concerning the properties and behavior of the target system (Levy, 2020). On this view, we should expect metaphors to increase understanding by telling us something about the world. For example, Richard Dawkins *represents* genes as selfish, which is (in some sense and in some contexts) a good representation, even if it is partially false.

To some extent, Levy and Camp take metaphors to be representational vehicles, with propositional content that is truth-evaluable. Maps have imagistic, spatial and descriptive representational content. But maps cannot be asserted. Metaphors can be asserted, and when they are, they can have propositional representational content.[2] For example, Levy refers to "metaphorical descriptions" (Levy, 2020, 291), as in the metaphorical description of information-transfer in biology. The metaphor *describes* biological processes as if they were transferring information. For Levy, metaphors are "explanatory vehicles," which are "the means by which information about explanatory facts is conveyed in the course of an explanatory episode" (Levy, 2020, 287).

Camp sometimes uses similar language. For example, "metaphors can enable speakers to communicate contents that cannot be stated in fully literal and explicit terms" (Camp, 2006, 1). "Because metaphorical utterances…express such complex contents in so few words, they are highly efficient vehicles for communication" (Camp, 2006, 3). Camp also refers to metaphors as devices for communication with representational content (Camp, 2007) and as representational vehicles (Camp, 2020, 307). This, therefore, is one option: metaphors represent a target as having certain features. This content is propositional and truth-evaluable when the metaphor is asserted. And when that content has the right good-making features and is connected in the right way to the understander, understanding is the result.

But there is another way to think about metaphors. This is to stress their invitational, imperative nature. For Walton, metaphors are invitations to play a game of make believe ("Come play this game!"), for Levy they are invitations to imagine ("Imagine it this way!"), and for Camp, they are invitations to adopt a certain perspective on a target ("Think about it this way!"). Semantically, invitations are more like imperatives

[2] Our notion of propositional representational content is meant to be broad enough to be consistent with many views in philosophy of science, mind, and language. The idea from philosophy of language that the representational content (of sentences) consists in truth-evaluable propositions goes back at least to Frege. In philosophy of science we have the view expressed by Frigg and Nguyen that the representational content of models is "the totality of what the representation says about the target," or the set of imputed features from the model to the target (Frigg & Nguyen, 2020, 182).

than indicatives, at least in the sense that they are non-assertoric and not truth-evaluable, and therefore do not have propositional representational content.[3]

Levy and Camp also express views that are in line with this second way of characterizing metaphors. For Camp, metaphors are "frames" which "guide our overall interpretation of a subject by providing a perspective" (Camp, 2020, 305). They "proffer a principle for organizing one's overall intuitive thinking about the target" (307). They "express" perspectives, which are "principles for interpretation rather than particular thoughts or contents in themselves" (308). "When a speaker utters a metaphor, she invites her hearer to use one characterization as a frame for structuring another…in uttering ["Juliet is the sun"], Romeo asks his hearers to use their (shared, contextually modulated) characterization of the sun to structure their characterization of Juliet" (Camp, 2017, 5).

Levy approves of Camp's view that metaphors are frames, and for Levy, a frame "directs one's thinking" and in so doing, "they allow one to utilize existing knowledge and reasoning skills. The understanding associated with metaphor therefore stems from the way in which it recruits pre-existing cognitive resources to new tasks and domains" (Levy, 2020, 293). For Levy, metaphorical utterances can guide our thinking, and we take advantage of this when we use them to frame an object in a new way. Thus, both Camp and Levy seem friendly to the view that metaphors can be thought of as invitations, uttered for their useful framing effects.[4]

There is an apparent tension: Either metaphors are asserted and have propositional representational content, or they are imperative invitations that are not asserted and do not have propositional representational content. We propose three ways to eliminate this tension. One is to characterize metaphorical utterances primarily as indicative representations with propositional content, and secondarily as invitations. This appears to be Levy's view. For Levy, the metaphor represents the world as being a certain way, often a way that would be false if taken literally. With metaphors, a scientist will offer such a representation with the secondary intention of inviting others to imaginatively explore the fictional world that it evokes. Here, the invitational nature of a metaphor follows from its assertoric representational nature.

Alternatively, we could characterize metaphorical utterances as primarily invitational, and secondarily representational (since we know that accepting the invitation will lead to a new characterization of the target, which results in new representations of the target). This occasionally seems to be Camp's view. For example, the speaker of a metaphor "expects her hearers to determine the content of her speech act *by way of*

[3] The utterance of a non-assertoric statement (e.g., an imperative or an interrogative) sometimes conveys information about how a speaker represents the world to be. For example, if I say "Close the door." this implies that I believe that the door is open. However, the imperative sentence itself does not say anything true or false, even for philosophers who develop purely truth functional theories of meaning (e.g., Davidson, 1979), and so, they do not have propositional representational content in the above sense. Sometimes non-assertorics have an "assertoric core", which is the part of their meaning that has representational content and which is modified by the mood indicator. But in other cases, non-assertorics do not have such a core (Ludwig, 1997). For example, Nike's slogan "Just do it" has the assertoric core: "do x," which is not a sentence with propositional representational content.

[4] The situation with metaphors in science is somewhat analogous to the situation concerning models: a mainstream view in philosophy of science is that models are representations. Another, sometimes overlapping view, is that models are invitations to imagine (Frigg, 2010; Frigg & Nguyen, 2016, 2019, 2020; Toon, 2012; Levy, 2015). The same two positions are held about metaphors. The question is whether metaphors are asserted representations, invitations, or both. We make the connection between models and metaphors more explicitly in section 6.4.

cultivating the relevant perspective: the perspective gives hearers the frame they need for thinking about the subject in order to identify how the speaker is claiming (or asking, or ordering) that subject to be" (Camp, 2007, 21–22, emphasis added). In more detail,

> In making a metaphorical utterance, a speaker asks her hearer to construe the subject *a* in terms of the frame *F*…The resulting reconfigured characterization of *a* is non-propositional, in the sense that it involves actually structuring one's intuitive thinking in the relevant pattern, and not just recognizing that such a pattern exists. But it also straightforwardly determines contents, including illocutionary contents. In the simplest cases…the metaphor's assertoric content is that *a* possesses those fitting properties that are most tightly matched to the most prominent and central fitting features of *F*. (Camp, 2017, 8)

A third possibility combines the above two: sometimes metaphors are best interpreted primarily as asserted indicatives and secondarily as imperatives, but sometimes it is the other way around. On this view, the very same metaphor might be best understood one way in one context (e.g., in a research paper) and another way in another context (e.g., in a classroom). That is, in one context, the metaphor is best understood as primarily indicative, and in the other, primarily imperative. In science, we should expect that many metaphors will be best understood as asserted indicatives, especially "dead" metaphors whose representational content has become conventional. We should also expect that some metaphors in science will be best understood as imperatives. This is because scientists know that metaphors structure thought, and they know that structuring thought in the right way can lead to progress. Therefore, it should come as no surprise if new metaphors are sometimes best understood primarily as proposals for changing how scientists structure their thought about a problem or phenomenon.

We adopt this third possibility, which flexibly combines the insights of assertoric and imperative views of metaphors. In every case, both aspects will be present and of epistemic interest. With respect to metaphors that increase understanding, we should note that a purely assertoric metaphor will only be as epistemically interesting as what it asserts. It might as well be literal. More interesting metaphors will be interesting because of their invitational nature. Still, a purely invitational metaphor (just in its role as an invitation) does not say anything about the world: for that, it needs its secondary, assertoric nature. We contend that it is the interaction between a metaphor's invitational and representational natures that makes it powerful, especially the power of the invitation to produce new representations. Because this interaction is (loosely speaking) causal, we take it to suggest a consequentialist epistemology.

4 Consequentialism and understanding

While some metaphors could increase understanding by introducing a single, epistemically good representation, the most impressive metaphors are epistemically valuable for what they *cause*, not (only) for what they assert. And what they cause is, in the best cases, new, better representations and new, better abilities. These two outcomes correspond to two ways that processing a metaphor can increase understanding, and

perhaps also to the two kinds of understanding found in Henk de Regt's influential account of scientific understanding. The first kind, for de Regt, is the kind of understanding we have when we are able to use a theory. He calls this "pragmatic understanding," and it is what renders a theory "intelligible" in his technical sense (De Regt, 2017, 91). The second is understanding a phenomenon, where that phenomenon is understood by a subject iff that subject possess an explanation of the phenomenon that is based on an intelligible theory, and conforms to the basic epistemic values of empirical adequacy and internal consistency (De Regt, 2017, 92). For de Regt, these two kinds of understanding are related in the sense that pragmatic understanding is necessary for explanatory understanding (De Regt, 2017, 92).

For de Regt, an agent has pragmatic understanding if they can recognize qualitatively characteristic consequences of a given theory without performing exact calculations (102). To increase pragmatic understanding, metaphors need to make a scientific theory easier to use, or provide the skills that a scientist requires for using a theory. A metaphor will cause scientists to apply a new perspective to aspects of a theory, which will create new dispositions. Insofar as these new dispositions allow the scientist to do something they could not do before, they can be understood as helping to increase that scientist's pragmatic understanding. The new dispositions can help a scientist to (recognize how to) construct models that bridge a theory to a target system, or construct or evaluate new explanations of a phenomenon. For example, the selfish gene metaphor prompts us to apply a new perspective to certain evolved behaviors, which makes certain useful inferences and predictions easier to make (e.g., about the behavior of genetically related individuals). Metaphors are helpful in getting people (especially students) to grasp and therefore be able to use theories, which at least partially explains why they are found so often in pedagogical contexts.[5]

Turning to explanatory understanding, we point out that explanations typically revolve around representations. Evaluating the quality of an explanation (on de Regt's account) will importantly consist in evaluating whether its constituent representations are empirically adequate and internally consistent. Thus, it is the quality of the representations that are (at least partially) responsible for the quality of an explanation. Metaphors can therefore aid in the search for explanatory understanding insofar as they cause us to form new representations that are, in some sense, better. But, what does it mean for a representation to be better, with respect to understanding? And in addition, returning to the first kind of understanding, how can we say when pragmatic understanding has been achieved? Our suggestion is consequentialist. While we will not offer any sustained argument for epistemological consequentialism (with respect to understanding) over its rivals, we take it to be motivated because its main rivals (deontology and virtue theory) do not seem prima facie plausible in this context. That is, we do not think understanding should be characterized wholly in terms of obeying epistemic norms of good reasoning, as in epistemic deontology. It is important to reason in a responsible way, but it is difficult to define understanding in those terms. While it makes sense to say that acting ethically is obeying moral norms, it makes less sense to say that understanding is obeying epistemic norms. Indeed, there does not seem to be any

[5] We do not limit pragmatic understanding to theories only: we allow for pragmatic understanding of models, concepts, equations, and perhaps material artefacts. But a staunch de Regtian can ignore this.

algorithm for generating understanding, let alone constituting it. With respect to virtue theory, understanding may often result from (or lead to) someone being an epistemically virtuous person. But we are not sure how to characterize scientific understanding itself in terms of virtue. There could of course be responses to these concerns, but since one cannot argue everything at once, for present purposes, one can just take epistemic consequentialism as a premise in what follows. That is, we follow the existing literature and portray understanding as a cognitive-epistemic *state*. Consequentialists, whether in ethics or epistemology, focus primarily on states as opposed to actions or agents, and they define right actions as those that promote good states of affairs.

Before moving forward, we want to emphasize that a consequentialist epistemology of metaphorical understanding goes beyond Camp's work on how we process metaphorical language by means of applying perspectives because Camp's account is "primarily descriptive" (Camp, 2020, 305). That is, her account is not meant to explain how metaphors provide understanding of target systems, but rather to illuminate "the cognitive structures and abilities that are generated by frames [like metaphors], and on the imaginative activities that exploit them" (ibid). She explains how we understand metaphors, and what metaphors do in reorganizing our thought. But this on its own does not distinguish between epistemically good and bad metaphors. Camp's account can explain what makes a metaphor more or less graspable, or cognitively powerful, but grasp is only part of understanding, because understanding is not only cognitive but also epistemic. We might have a very powerful, graspable metaphor that nevertheless decreases understanding, and a complete epistemology of metaphor must have the resources to explain what is wrong with such a metaphor.[6]

For example, Kampourakis (2016) argues that geneticists should not refer to the genome as the "book of life" because this oversimplifies the complexity of genetics and introduces a false sense of completeness concerning our scientific understanding of genes. Camp's view nicely explains what happens to a non-scientist when they are exposed to this metaphor in terms of bringing a new perspective to bear on a topic which alters their characterization of genetics. Despite grasping the metaphor successfully, the metaphor goes wrong in its effects on the agent's representations and abilities.[7]

So, what is the state of understanding that good metaphorical utterances promote? Our preferred account of understanding captures the intuitions motivating both pragmatic and explanatory understanding. This is Daniel Wilkenfeld's "Understanding as Compression" (UC) account. Here is the official statement of UC, including explanatory footnotes:

[6] Of course, Camp's account is very good as an account of what has been called linguistic or conceptual understanding. Traditionally, these kinds of understanding are not discussed in philosophy of science. We think this is misguided: conceptual understanding is important and should count as a kind of understanding proper. For the rest of this paper, however, we will focus on the more traditional notion of understanding of a scientific system of interest, rather than a concept.

[7] For a more extreme example, consider the nineteenth century metaphorical likening of women and the so-called "lower races" in terms of cognitive abilities and brain size/structure (Stepan, 1986). We want to be able to explain not just what made this metaphor graspable for the scientists at the time, but why it produced *mis*understanding about both gender and race.

> Understanding as Compression (UC): A person p_1 understands object o in context C more than another person p_2 in C to the extent that p_1 has a representation/process pair that can generate more information of a *kind* that is useful[8] in C about o (including at least some higher order information about which information is relevant in C)[9] from an accurate, more minimal representation.[10] (Wilkenfeld, 2019, p. 4)

Understanding, on this picture, involves both what we think of (the content of the representations) and how we think about it (the process by which the representation is decompressed to yield useful information).[11] It allows us to analyze the extent to which a metaphor increases pragmatic understanding in the sense that it focuses on changes with respect to the usefulness of the representation/process pair in terms of the abilities of the agent. It allows us to analyze the extent to which a metaphor increases explanatory understanding in the sense that it focuses on changes with respect to the representations involved in terms of their accuracy and minimality.

The general idea behind UC is that what understanding does for us is enable us to get by with a lighter cognitive load. Metaphors increase understanding by changing our cognitive situation so that it is possible, for a given target in a given context, to do more with less. This account both explains what understanding is, and (indirectly) why it is valuable. It is valuable because it allows us to process information more efficiently, and helps us reconstruct information from only partial data/memories. And because compressing representations involves exploiting genuine patterns in the data, it allows us to avoid over-fitting our predictions to outlying data-points (Wilkenfeld, 2019). These cognitive benefits are in addition to the immediate phenomenological payoff (or "aha!" feeling) discussed in Gopnik (1998).

It is also worth flagging that the mere fact that representations are added to a system does not entail that it is thereby less compressed. For example, if one had an extremely lengthy encoded sequence that went "01010101010101…" adding the representation *that the sequence has alternating 0s and 1s* actually enables one to compress the system overall.

It is worth noting that one advantage of UC is that it aspires to go beyond scientific understanding, so if we can show how metaphors play a role in the kind of understanding characterized by UC, there will be fairly broad applicability. Given the focus of Camp and others on examples of metaphors from poetry and fiction, this benefit

[8] The usefulness of information is determined relative to the agent's aims, which can range from unlocking the secrets of nature to passing an exam or getting to work on time.

[9] This clause is required to avoid the counterintuitive result that someone who can generate more of the right kind of information but cannot discern which information is relevant understands better than someone with more targeted understanding.

[10] Strictly speaking, there should also be a provision that the representation does not generate too much false information, lest one think that a contradiction plus classical logic was maximally understanding-producing. We doubt most philosophers (and certainly most non-philosophers) are worried about the implications of logical explosion for an account of understanding, so we leave this condition implied for the remainder of the paper.

[11] Levy provides an account that is similar to an earlier version of this account (specifically, the account found in Wilkenfeld, 2013). What we say in the following is consistent with the main thrust of Wilkenfeld's earlier account, as well as Levy's, namely, that metaphors enable understanding by increasing the number or quality of inferences a scientist can make after being exposed to the metaphor. However, we take Wilkenfeld's more recent account to allow for finer distinctions among degrees of understanding.

might help tie together scientific metaphors with what has already been said elsewhere. A second advantage is that UC allows for the in-principle quantification of understanding: an agent understands more insofar as they can generate more useful information about an object in a context from more accurate, more minimal representations.

Now, our suggestion is to define good metaphors as those that promote understanding in the sense of UC. In other words, A person at t_2 understands object o in context C more than the same person at t_1 in C to the extent that at t_2 the person has a representation/process pair that can generate more information of a kind that is useful in C about o. If a person wants to understand o, and there is an action that will get them from the state at t_1 to the state at t_2, ceteris paribus, they should perform that action. Metaphors can now be evaluated relative to their consequences on mental representations and processes that are relevant for understanding. Importantly, metaphors are not (only) to be evaluated with respect to their quality as asserted propositional representations, but (also) with respect to their consequences (as invitations) on the dispositions of the agent.

In more detail, UC allows us to be specific about the two ways that metaphors increase understanding: by having good downstream consequences on a) our cognitive processes, or b) our representations. One way to understand the former is to think about skills. Most philosophers who disagree about the nature of understanding still agree that improving our skills, including conceptual skills and inferential skills, can enhance understanding (De Regt, 2017; Elgin, 2017; Khalifa, 2017). Processing a metaphor can enhance understanding by leading to an increase in the quality of the connection between the agent and their set of representations about an object. Agents must not merely have access to good representations, however, they must be able to use them. Following the instructions of the metaphor can cause the agent to produce the useful representations on their own. The agent can then recall the useful (set of) representations easily, by recalling the metaphor. Also, the metaphor causes an entire perspective to be applied, which guides the agent in making a potentially infinite number of inferences about the target. The perspective is understanding-conducive to the extent that it enables the agent to infer more useful information from a representation in a context. For example, an agent might have accurate representations of the economic trends in a particular market, but they might not yet easily recognize qualitatively characteristic consequences of events taking place in that (or other) markets. Their understanding is still at a low level. However, the imaginative actions they perform when exposed to, say, Adam Smith's metaphor of the invisible hand, can help in developing the relevant abilities. And all of this happens efficiently, because the simple mention of the metaphor can cause the agent to bring an entire set of dispositions and resultant mental representations to bear in a way that helps to draw out useful information about economic markets.

The second way for metaphors to increase understanding concerns the set of representations possessed by the agent. Metaphors that have representational content might cause an agent to obtain a single new, useful representation. But metaphors are also able to transform an agent's cognitive approach by producing an open-ended set of new and potentially useful representations. The set of representations produced by a metaphor can be more minimal and more accurate than an existing set, even if the metaphor itself has false or inaccurate propositional representational content (when uttered as an assertion). And the more minimal and accurate the resulting

representations are, the more understanding is increased. Processing a metaphor affects our set of mental representations by changing which features of that target matter most, in terms of prominence, intensity, and centrality. Good metaphors are those that tend to do this in a way that improves our set of representations of a target, in the sense that the resulting representations are either more accurate or more minimal (or both). For example, the metaphor of genes as selfish agents causes us to alter our mental representation of evolutionary genetics in a way that increases its accuracy and minimality. However, this only works insofar as genes actually do have the features that the selfish gene metaphor inspires us to represent them as having.

We now apply this consequentialist application of the UC model to some examples.

5 Examples

5.1 The market is governed by an invisible hand

Adam Smith (1776/1963) argued that prices fluctuating naturally was better for individuals in a market than any intentional, system-level oversight. Supply and demand on their own lead to resources ending up where they are most efficiently used, without needing anyone to explicitly guide the process, e.g., through redistribution. Smith summarized his thought in a slogan: the market is guided by an invisible hand. We can think of this metaphor as having representational content, by representing markets as described above. But that representation on its own does not exhaust what is interesting about it. What makes this metaphor so useful is the way it enables us to very easily generate sets of representations of particular markets that are both minimal and accurate, and to draw information about real markets from those representations. Looking upon the actions of the market as a whole from the outside, we notice that certain parameters will appear to shift and sway *as if* guided, eventually but certainly, towards more efficient arrangements. It is much easier to predict what such a hand will do (the most economical, rational thing) than to predict the decisions of hundreds, thousands, millions, or billions of individuals. And thinking about resources being moved by a hand is a very simple, powerful way to remember and use that idea.

5.2 Force is unbalanced weight

At Galileo's time, Aristotelians explained the cause of motion in terms of the natural tendency of elements to return to their "natural" places in a spherical universe. The concentric sphere that was the natural home of fire contained the natural home of air, which contained that of water, which contained that of earth. And the natural home of earth was at the center of the universe. "Earthy" rocks will fall through the air and sink through water, trying to get to the center of the universe, while bubbles of air will rise through water to get above it. Everyday objects were of mixed composition, so their "natural" motion could be understood via an analysis of the ratios of the elements of which they were composed.

Galileo tried to shift the central metaphor about the cause of motion in physics away from teleological strivings for natural places. In Chapter 6 of *De Motu* he claimed that motion is caused by *forces*, which he explained with the metaphor of unbalanced

weight (Machamer & Woody, 1994, 216). Thus, "A heavy body tends to move downward with as much force as necessary to lift it up, i.e., it tends to move downward with the same force with which it resists rising" (Galilei, 1960, 54). This metaphor of force as unbalanced weight became the foundation for Galileo's work on inclined planes and hydrostatics. The balance was a good source for the metaphor, being a millennia-old device commonly used to weigh items and do accounting. Galileo tended to consider motions taking place in the same medium, which helped to reduce complicated problems to simpler ones where the main or only variable was weight, a development so useful that Peter Machamer and Andrea Woody call the balance metaphor an "understanding device" (Machamer & Woody, 1994, 216). It seems plausible that the force-as-unbalanced-weight metaphor is best interpreted as primarily a proposal for how to think about force rather than a proposal about what force "really is."

The metaphor of motion as unbalanced weight unifies diverse phenomena and allows us to account for many kinds of motion: in other words, it allows us to do more with less. And the metaphor provides visualizable solutions to problems. These visualizations make possible the geometrical representations of problems that modern students of physics still employ in statics with force vector diagrams. Machamer and Woody claim that visualizability is not necessary for understanding, but it is "important for teaching beginning students and will aid their ability to understand" (Machamer & Woody, 1994, 220; see also De Haro & De Regt, 2018; De Regt, 2014).

In processing the force-as-unbalanced-weight metaphor, we create new a new kind of representation of individual motions (and motion in general), especially through the use of visualization, spatial information, and previous dispositions concerning balanced and unbalanced weight. Meanwhile, we also gain new skills that enable us to generate useful information from these representations, for example, about forces experienced by objects and their resultant motions. These consequences of the metaphor explain why it has persisted so long.

5.3 Genes are selfish

Richard Dawkins (2016) famously explains evolution as the product of genes wanting to reproduce themselves and spread as much as possible. The gene, for Dawkins, is the central object on which natural selection acts, as opposed to the species or organism. This metaphor has been important for evolutionary biology, although it continues to be criticized. Caricatures of Dawkins's view portray his metaphor as a false description of genes. That it may be. However, the success of the metaphor is easy to explain if we understand Dawkins as defending the usefulness of a particular scientific perspective, perhaps because it *leads to* better representations which are themselves accurate, minimal, and useful for making contextually fruitful inferences about evolved traits.

Still, we do not want to claim that the metaphor is perfect. Interestingly, if the metaphor is flawed, our account provides a new explanation for why. Rather than claiming that the metaphor is flawed because it is false or inaccurate, we claim that utterances of the metaphor are flawed for their (foreseeable) negative consequences. This is generalizable. Characterizing the failure of a metaphor only in terms of its

failure as a representation misses some other dangerous ways that a metaphor can fail. A racist metaphor that is used by a politician or scientist might be epistemologically poor because it is a bad representation. But it should also be criticized for the epistemologically bad consequences it has on those who process it. This is perhaps the main reason we must be vigilant with metaphors: because they lay the groundwork for new theories and models, which inherit the biases contained in their seed metaphors. Once theories and models based on a metaphor are fully developed, these biases become hard to detect and eliminate. And this is true even at the justification stage, because biases can be built into the justification process itself (Dellsén, 2019).

5.4 The brain codes and decodes information

As the wires of a telegraph carry signals that can be decoded, pathways in the brain carry signals from the senses, which can be "decoded" into information about the environment. This metaphor probably goes back to the work of physiologist Edgar Adrian, though it appears in a crude form already in Hermann Helmholtz and Emil du Bois-Reymond (Garson, 2019). The idea is that we can think about the codes carried by neurons as containing information which can be decoded: there is a sender and a receiver, and the signal has some causal effect on the receiver (Brette, 2019).

The coding metaphor allowed for the formulation of new questions that moved researchers away from the mechanistic details of the action potential, enabling more abstract and teleological thinking about the relationship between neural stimuli and responses (Garson, 2019). But some now argue that it has outlasted its usefulness. The problem, according to Roman Brette, is that the notion of "information" at work here is inappropriate for investigating what neurons do. The sense of information has typically been Shannon information, but what we really need is something more like the rich kind of information carried in (spatial and temporal) models of the environment. Additionally, the metaphor encourages scientists to characterize the causal structure of the brain incorrectly: there are no linear, algorithmically decodable signals. Instead, we have dynamic, circular and distributed processes (Brette, 2019).

On our account, the problem with this metaphor is not that it represents falsely, or not sufficiently accurately. Instead, we want to say that the metaphor initially had understanding-conducive consequences. But now, in a context where we have much more information about the brain, the consequences of the metaphor are not as understanding-conducive. This is possible for any metaphor, which might have good consequences in one context, and bad consequences in another. Additionally, a metaphor can have *initial* consequences that are good, but long-term consequences that are bad. This at least partially explains the reactions scientists have to the metaphor now. Some argue that the metaphor could be saved by becoming more specific to reflect our modern understanding of the brain (Birch & Smortchkova, 2019), while others claim it should be replaced with a new, different metaphor that better reflects the current context (Deacon & Rączaszek-Leonardi, 2019). In both cases, the claim is motivated by how we might improve the effects of the metaphor in guiding working scientists or educating future students.

In sum, a metaphor can increase the understanding of a person about an object in a context to the extent that the application of that metaphor leads to the creation of more accurate, more minimal mental representations, or by causing the agent to gain new (or recruit existing) abilities to generate more information that is useful in the context about the object (including at least some higher order information about which information is relevant in the context). And of course, a good metaphor could be good in the sense that it caused both of these things to happen, instead of just one. And, as with ethical consequentialism, we can consider the effects of a metaphor both for an individual or a group. For example, Galileo's force-as-unbalanced-weight metaphor was good for Galileo, not so good for his academic enemies, and very good for generations of scientists who followed, as it provided an easy way to generate representations of motion that are easy to use.

Having presented our account of metaphorical understanding, we will now explore some of its applications, and see how it intervenes in existing debates in philosophy of science.

6 Consequences

6.1 Discovery, creativity, and praiseworthiness

According to Berys Gaut, "Metaphor-making is a paradigm of the creative use of imagination" (Gaut, 2003, see also Veit & Ney, 2021). Dramatic advances in scientific understanding have been the result of metaphors. Indeed, "it is often the idea of a new metaphor that guides the scientist. An object, an event, is suddenly perceived in an unusual and revealing light, as if someone abruptly tore off a veil that, till then, had covered our eyes" (Jacob, 2001, 119). Francis Jacob cites several examples of such sudden leaps forward in understanding: Newton seeing the moon as a ball thrown around the world, Plank seeing the radiation of heat as a "hail of quanta," and William Harvey seeing the heart as a mechanical pump. These metaphors cause new mental representations (or alter old ones), many of which are more accurate and more useful than what came before. Citing similar examples, Gaut claims, "science often spins its theories from a metaphorical source" (Gaut, 2003, 287). How does this happen? It is true that scientists sometimes speak as if imagination itself, a force Shakespeare described as the "muse of fire," whisked them away, willing or not, towards a major discovery via metaphor. But if that were the case, the scientist (as a mere medium) wouldn't deserve any praise for their discovery: the muse of fire would.

The present account of metaphorical understanding helps to address this tension by reminding us of two important properties of perspectives: First, they are open-ended sets of dispositions, which are nevertheless *tightly connected*. In other words, we bring *entire* perspectives to bear on a topic, with little effort, despite the fact that the perspective contains a potentially infinite number of new dispositions and could result in a potentially infinite number of new representations. A scientist might intend to bring a new perspective to bear on a topic, anticipating its good consequences, and that is certainly praiseworthy insofar as their guess is warranted and/or the metaphor actually leads to new understanding. They might not know in advance how "transformational" a particular metaphor will be, but they might still

have very good reasons for thinking that a particular perspective will be understanding-conducive.[12] That said, the scientist may still be surprised by what happens when the full perspective is brought to bear: how quickly things fall into place, and how different the target now looks, as the rest of the perspective snaps into place. This would explain the phenomenology of sudden understanding that scientists experience in moments of discovery via metaphor.

6.2 Metaphorical understanding and factivity

It is often said that metaphors are literally false, and this has been significant for motivating the question of how metaphors can have epistemic value. For instance, Levy claims that less accurate metaphors can be more understanding-conducive than more accurate ones, which naturally prompts the question: how can we have epistemic success via something that is false? This question relates to the "factivity" of understanding. It is (relatively) uncontroversial that *knowledge* is factive in the sense that if S knows that p, p must be true. But it is unclear whether understanding is factive. One might think that metaphors support the claim that understanding is not factive, because metaphors look like false representations that nevertheless form part of the content of scientific understanding.

There are a number of ways to respond. One is to try to separate the accurate from the inaccurate parts of a representation. For example, Insa Lawler (2019) argues that the false parts of a "felicitous falsehood" merely *enable* understanding, but do not form part of the content of the understanding: only the parts of the representation that are accurate make up the content of understanding. Stephen Yablo (2020) presents a similar argument: scientific representations only assert the things which hold of their targets. They do not assert mathematical, idealized, abstract, or metaphorical aspects of their content. Opponents of this strategy worry whether the content of what is asserted can really be decomposed so easily (see, e.g., Rice, 2019).

A second line of response would be to defend the existence of something like *metaphorical truth*. Even if a metaphor does not literally represent its target system accurately, it can *metaphorically* represent it accurately, perhaps by metaphorically instantiating or exemplifying key features of the actual target system. Thus, Catherine Elgin claims that ballet sometimes "literally exemplifies properties such as grace, delicacy, and beauty; and metaphorically exemplifies properties such as love and longing, weightlessness and ethereality" (Elgin, 2017). We might come to understand something about love and longing from a ballet, even if love and longing are not literally exemplified, by being exposed to the love and longing that is *metaphorically* (but, still, *actually*) exemplified by the dance.

We prefer to refocus the discussion from the content of metaphors to their consequences. The real question, in our opinion, concerns the factivity of the representations *produced* by metaphors. We can imagine an uttered metaphor leading an agent to

[12] Margaret Boden calls moments of transformational creativity "the deepest cases of creativity." These involve "someone's thinking something which, with respect to the conceptual spaces in their minds, they couldn't have thought before. The supposedly impossible idea can come about only if the creator changes the pre-existing style in some way. It must be tweaked, or even radically transformed, so that thoughts are now possible which previously (within the untransformed space) were literally inconceivable" (Boden, 2003, 6). We think the metaphors listed above should count as transformative, and we claim our account explains why: new perspectives can radically change existing conceptual spaces.

produce a set of literal (or dead-metaphorical) representations that are very accurate and very minimal, and also new abilities that assist in tracking the truth. We can also imagine a metaphor causing an agent to produce a set of representations that are inaccurate but useful, as well as useful abilities that do not track the truth. The question about factivity concerns whether, when, and why the latter sort of consequences should count as understanding. And this question can be answered without appeal to the content of the metaphors that produced such consequences. Thus, we claim that the propositional representational content of metaphors are not immediately relevant to questions of factivity.[13]

6.3 Special skills and testimony

Another important issue in the recent literature on understanding concerns the relationship between understanding and skills. Some argue that certain skills are partially constitutive of understanding, and this is what separates understanding from knowledge. Kareem Khalifa argues against this position, claiming that understanding requires no *special* skills, that is, no skills that aren't also required for possessing knowledge (Khalifa, 2017, 51–79). One lightening rod for this issue has been the possibility of testimonial understanding (Hazlett, 2022; Hills, 2016). If a reliable source informs us of an adequate explanation of a phenomenon, do we thereby understand the phenomenon, or must we first develop some special skills concerning that phenomenon (and/or its explanation) to be able to say that we *really* understand it?

A consequentialist account of metaphorical understanding suggests a graded answer to this question. When exposed to certain easy-to-process metaphors, it is straightforward to apply the new dispositions that come with the new perspective, and generate new representations that may or may not be useful. New abilities might appear almost automatically upon exposure to the metaphor. Some scientific metaphors are like this: upon exposure to the suggestion that one phenomenon can be treated as another, scientists are immediately able to apply the (e.g., mathematical) abilities to the former which were developed and previously reserved for the latter. Because the good consequences follow directly from exposure to the metaphor, we can attribute epistemic value to the metaphorical utterance without reference to any special skills of the listener.

But some metaphors pay large dividends for time invested. The force-as-unbalanced-weight metaphor, for example, has given generations of scientists new skills for drawing useful information about force and motion, but the relevant representations and abilities come only through practice, which mere exposure does not provide. This metaphor increases understanding because of the impact it has on the minds of scientists, and those good consequences cannot be wholly attributed to the utterance of the metaphor itself. The effort and skills of the audience matter.

Thus, with respect to the possibility of understanding via testimony, we claim that some metaphors can increase understanding via mere utterance and comprehension. Of course, the metaphor has to be processed, but in these cases, the processing is no different than the usual kind of processing that is involved in literal speech comprehension. In other cases, additional understanding is produced when new cognitive

[13] We think it is plausible that the strategy here could be extended to other tools of reasoning that are discussed with respect to factivity, including models, diagrams, analogies, and so on.

abilities are forged through cognitive effort, and in these cases, more than mere linguistic comprehension is necessary. These extra skills and effort might very well go beyond what is required for grasping propositional knowledge.

6.4 Metaphors and models

Both Camp and Levy claim that metaphors involve a kind of surrogative reasoning that is in the same family as reasoning with models. Roughly, models say something very specific about the target system, while metaphors are less specific. This raises the following question for our view: do we extend it to models, or not? That is, do we advocate focusing on the consequences of adopting a model, rather than on its asserted propositional representational content?

The idea that models directly assert something about the world is widespread. Indeed, this is perhaps the primary function of models when characterized as "mediators" between theory and reality: they enable a theory (whose content is a set of models, not claims about the world) to make specific predictions about target systems. Standard examples of representational models include "the billiard ball model of a gas, the Bohr model of the atom, the Lotka-Volterra model of predator-prey interaction, the Mundell-Fleming model of an open economy, and the scale model of a bridge" (Hartmann & Frigg, 2012). As Ronald Giere puts it, models are representations in "one of the more general senses now current in cognitive psychology" (Giere, 1988, 80; see also Morgan & Morrison, 1999, 5). These models might refer by fiat, or by isomorphism, partial isomorphism, or some other kind of similarity relation. "If there is any theme that unites philosophers as regards models, it is that of representation: models are generally taken as representations" (Knuuttila & Boon, 2011, 309).

However, some philosophers argue that models should not be thought of as representations. Anna Alexandrova is one example, who portrays models as tools used to generate testable empirical hypotheses (Alexandrova, 2008). Tarja Knuuttila and Mieke Boon argue that models are not primarily representational, but artefactual: they are tools to do epistemic things (Knuuttila & Boon, 2011; see also Knuuttila, 2011, 2021).

We want to take a different path, by focusing on the relationship between the assertoric and imperative natures of models. Perhaps some models are best thought of as assertions of (sets of) propositions, and in considering them, we are invited (secondarily) to explore the idea they suggest. In other cases, models are best thought of as invitations (to imagine or (re)conceptualize a target system), which, if accepted, will lead to new representations and abilities. In either case, the epistemic value of the model with respect to understanding will depend on the downstream consequences of embracing the model. We do think this could be a helpful view, even if there are important differences between models and metaphors at the semantic or functional level.[14]

[14] We should note a possible objection. If the epistemic value of a metaphor or a model is based on its consequences, this might imply that we cannot say whether a model or metaphor is epistemically good or bad until we have seen it play out for some time. However, a metaphor or model can be judged as epistemically good or bad based on its *anticipated* consequences. And when these consequences are not possible to foresee, scientists should refrain from epistemic judgment and try the model out for themselves. We think this coheres with actual scientific practice. For example, scientists often judge a model to be good or bad, only later to reverse their decision. This again makes sense on a consequentialist epistemology.

7 Conclusion

Metaphors can be evaluated as understanding-conducive in terms of their effects on the cognitive-epistemic states of agents. In this paper, we developed a consequentialist position that defines epistemically desirable states of affairs in terms of Wilkenfeld's Understanding as Compression account, and defined good metaphors as those that promote such states. We identified the state of understanding in terms of representations and abilities, and discussed different ways that metaphors could bring desirable changes about. Finally, we explored some of the different ways that the presented account interacts with broader issues in philosophy of science, including discovery, creativity, factivity, testimony, skills, and reasoning with models.

One avenue for future research would be to see whether any special difficulties arise when attempting to extend this account to scientific understanding via analogies, visualizations, just-so-stories, thought experiments, and particular models. A second avenue would be to look more carefully at the neuroscience of metaphor-use and understanding acquisition, to see what can be learned from the physical mechanisms involved in one or both processes.

Acknowledgements The research was supported by the Swiss National Science Foundation (grant number PZ00P1_179986), the University of Tübingen, and the University of Pittsburgh. We would also like to thank three anonymous referees for helpful feedback, as well as audiences at the Erasmus University Rotterdam.

Authors' contributions The authors contributed equally to the paper. Author names are listed in alphabetical order.

Declarations

Conflicts of interest/competing interests There are no conflicts of interest.

References

Alexandrova, A. (2008). Making models count. *Philosophy of Science, 75*, 383–404.

Aubusson, P. J., Harrison, A. G., & Ritchie, S. M. (2006). *Metaphor and analogy in science education*. Springer.

Beger, A., & Jäkel, O. (2015). The cognitive role of metaphor in teaching science: Examples from physics, chemistry, biology, psychology and philosophy. *Philosophical Inquiries, 3*(March), 89–112.

Birch, J., & Smortchkova, J. (2019). From the 'coding metaphor' to a theory of representation. *Behavioral and Brain Sciences, 42*, e220. https://doi.org/10.1017/S0140525X19001456

Black, M. (1962). *Models and metaphors*. Ithaca and London: Cornell University Press.

Boden, Margaret A. 2003. *The Creative Mind: Myths and Mechanisms*. 2nd edition. London ; New York: Routledge.
Brette, R. (2019). Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences, 42*, e215. https://doi.org/10.1017/S0140525X19000049
Camp, E. (2006). Metaphor and That Certain 'Je Ne Sais Quoi'. *Philosophical Studies, 129*, 1–25.
Camp, E. (2007). Showing, telling and seeing. Metaphor and 'poetic' language. *Baltic International Yearbook of Cognition, Logic and Communication, 3*(1). https://doi.org/10.4148/biyclc.v3i0.20
Camp, E. (2009). Two varieties of literary imagination: Metaphor, fiction, and thought experiments. *Midwest Studies in Philosophy, 33*(1), 107–130.
Camp, E. (2017). Why metaphors make good insults: Perspectives, presupposition, and pragmatics. *Philosophical Studies, 174*(1), 47–64. https://doi.org/10.1007/s11098-015-0525-y
Camp, E. (2020). Imaginative frames for scientific inquiry: Metaphors, telling facts, and just-so stories. In A. Levy & P. Godfrey-Smith (Eds.), *The scientific imagination* (pp. 304–336). Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190212308.001.0001/oso-9780190212308-chapter-14
Davidson, D. 1979. "Moods and Performances." In *Meaning and Use*, edited by A. Margalit, 9–20. Reidel.
Dawkins, R. (2016). *The extended selfish Gene*. Oxford University Press.
De Haro, S., & De Regt, H. (2018). A precipice below which lies absurdity? Theories without a Spacetime and scientific understanding. *Synthese*. https://doi.org/10.1007/s11229-018-1874-9
De Regt, H. (2014). Visualization as a tool for understanding. *Perspectives on Science, 22*(3), 377–396. https://doi.org/10.1162/POSC_a_00139
De Regt, H. (2017). *Understanding scientific understanding*. Oxford University Press.
Deacon, T. W., & Rączaszek-Leonardi, J. (2019). Abandoning the code metaphor is compatible with semiotic process. *Behavioral and Brain Sciences, 42*, e224. https://doi.org/10.1017/S0140525X19001419
Dellsén, F. (2019). The epistemic impact of theorizing: Generation bias implies evaluation bias. *Philosophical Studies, December*. https://doi.org/10.1007/s11098-019-01387-w
Dellsén, F. (2020). Beyond explanation: Understanding as dependency modelling. *The British Journal for the Philosophy of Science, 71*(4), 1261–1286. https://doi.org/10.1093/bjps/axy058
Elgin, C. Z. (2017). *True enough*. MIT Press.
Frigg, R. (2010). Models and fiction. *Synthese, 172*(2), 251–268.
Frigg, R., & Nguyen, J. (2016). The fiction view of models reloaded. *The Monist, 99*(3), 225–242.
Frigg, R., & Nguyen, J. (2019). Mirrors without warnings. *Synthese, May., 198*, 2427–2447. https://doi.org/10.1007/s11229-019-02222-9
Frigg, R., & Nguyen, J. (2020). *Modelling*. : *An Opinionated Introduction to Scientific Representation*. Synthese Library. Springer International Publishing. https://doi.org/10.1007/978-3-030-45153-0.
Galilei, G. (1960). *On motion, and on mechanics: Comprising De Motu (ca. 1590) translated with introduction and notes by I.E. Drabkin, and Le Meccaniche (ca. 1600) translated with introduction and notes by Stillman Drake*. University of Wisconsin Publications in medieval science 5. University of Wisconsin Press.
Garson, J. (2019). The origin of the coding metaphor in neuroscience. *Behavioral and Brain Sciences, 42*, e227. https://doi.org/10.1017/S0140525X19001316
Gaut, B. (2003). Creativity and imagination. In B. Gaut & P. Livingston (Eds.), *The creation of art: New essays in philosophical aesthetics* (pp. 148–173). Cambridge University Press https://risweb.st-andrews.ac.uk/portal/en/researchoutput/creativity-and-imagination(0769616d-8027-4b94-8473-315ecc2bb954).html
Giere, R. (1988). *Explaining science: A cognitive approach*. University of Chicago Press.
Gopnik, A. (1998). Explanation as orgasm. *Minds and Machines, 8*(1), 101–118.
Hartmann, S., & Frigg, R. (2012). Models in science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/models-science
Hazlett, A. (2022). Understanding and testimony. In J. Lackey & A. McGlynn (Eds.), *The Oxford handbook of social epistemology*. Oxford (Forthcoming).
Herrington, E., & Jablonka, E. (2020). Creating a 'gestalt shift' in evolutionary science: Roles for metaphor in the conceptual landscape of the extended evolutionary synthesis (EES). *Interdisciplinary Science Reviews, 45*(3), 360–379. https://doi.org/10.1080/03080188.2020.1794383
Hills, A. (2016). Understanding why. *Noûs, 50*(4), 661–688.
Jacob, F. (2001). Imagination in art and science. *Kenyon Review, 23*, 113–121.
Kampourakis, K. (2016). The bad use of metaphors and the use of bad metaphors. *Science & Education, 25*(9–10), 947–949. https://doi.org/10.1007/s11191-016-9870-2
Kendall-Taylor, N., Erard, M., & Haydon, A. (2013). The use of metaphor as a science communication tool: Air traffic control for your brain. *Journal of Applied Communication Research, 41*(4), 412–433. https://doi.org/10.1080/00909882.2013.836678

Khalifa, K. (2017). *Understanding, explanation, and scientific knowledge*. New York: Cambridge University Press.
Kind, A. (2021). Can imagination be unconscious? *Synthese, August., 199*, 13121–13141. https://doi.org/10.1007/s11229-021-03369-0
Knuuttila, T. (2011). Modelling and representing: An artefactual approach to model-based representation. *Studies in History and Philosophy of Science Part A, Model-Based Representation in Scientific Practice, 42*(2), 262–271. https://doi.org/10.1016/j.shpsa.2010.11.034
Knuuttila, T. (2021). Epistemic artifacts and the modal dimension of modeling. *European Journal for Philosophy of Science, 11*(3), 65. https://doi.org/10.1007/s13194-021-00374-5
Knuuttila, T., & Boon, M. (2011). How do models give us knowledge? The case of Carnot's ideal heat engine. *European Journal for Philosophy of Science, 1*(3), 309–334. https://doi.org/10.1007/s13194-011-0029-3
Lawler, I. (2019). Scientific understanding and felicitous legitimate falsehoods. *Synthese, December., 198*, 6859–6887. https://doi.org/10.1007/s11229-019-02495-0
Levy, A. (2015). Modeling without models. *Philosophical Studies, 172*(3), 781–798. https://doi.org/10.1007/s11098-014-0333-9
Levy, A. (2020). Metaphor and scientific explanation. In A. Levy & P. Godfrey-Smith (Eds.), *The scientific imagination* (pp. 280–303). Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190212308.001.0001/oso-9780190212308-chapter-13
Ludwig, K. (1997). The truth about moods. *ProtoSociology, 10*, 19–66. https://doi.org/10.5840/protosociology1997102
Machamer, P., & Woody, A. (1994). A model of intelligibility in science: Using Galileo's balance as a model for understanding the motion of bodies. *Science and Education, 3*, 215–244.
Morgan, M., & Morrison, M. (1999). *Models as mediators*. Cambridge University Press.
Nersessian, N. (1984). *Faraday to*. Einstein: *Constructing Meaning in Scientific Theories*. Science and Philosophy. Springer Netherlands. https://doi.org/10.1007/978-94-009-6187-6.
Nersessian, N. J. (2015). The cognitive work of metaphor and analogy in scientific practice. *Philosophical Inquiries, 3*(1), 133–156. https://doi.org/10.4454/philinq.v3i1.118
Nyrup, R. (2018). *Of water drops and atomic nuclei: Analogies and pursuit worthiness in science*. June: The British Journal for the Philosophy of Science. https://doi.org/10.1093/bjps/axy036.
Rice, C. (2019). Models Don't decompose that way: A holistic view of idealized models. *The British Journal for the Philosophy of Science, 70*(1), 179–208. https://doi.org/10.1093/bjps/axx045
Smith, A. (1963). *An inquiry into the nature and causes of the wealth of nations*. Edited by Dugald Stewart. Vol. 1. Wiley online library.
Spranzi, M. (2004). Galileo and the mountains of the moon: Analogical reasoning, models and metaphors in scientific discovery. *Journal of Cognition and Culture, 4*(3–4), 451–483. https://doi.org/10.1163/1568537042484904
Stepan, N. L. (1986). Race and gender: The role of analogy in science. *Isis, 77*(2), 261–277. https://doi.org/10.1086/354130
Stuart, M. T. (2016). Taming theory with thought experiments: Understanding and scientific Progress. *Studies in History and Philosophy of Science Part A, 58*(August), 24–33. https://doi.org/10.1016/j.shpsa.2016.04.002
Stuart, M. T. (2018). How thought experiments increase understanding. In M. T. Stuart, Y. Fehige, & J. R. Brown (Eds.), *The Routledge companion to thought experiments* (pp. 526–544). Routledge. https://www.routledgehandbooks.com/doi/10.4324/9781315175027.ch30
Stuart, M. T. (2019). Towards a dual process epistemology of imagination. *Synthese., 198*, 1329–1350. https://doi.org/10.1007/s11229-019-02116-w
Toon, A. (2012). *Models as*. Make-Believe: *Imagination, Fiction and Scientific Representation*. New Directions in the Philosophy of Science. Palgrave Macmillan UK. https://doi.org/10.1057/9781137292230.
Veit, W., & Ney, M. (2021). *Metaphors in Arts and Science*. https://doi.org/10.13140/RG.2.2.23949.15841
Walton, K. L. (1993). Metaphor and prop oriented make-believe. *European Journal of Philosophy, 1*(1), 39–57. https://doi.org/10.1111/j.1468-0378.1993.tb00023.x
Wilkenfeld, D. A. (2013). Understanding as representation manipulability. *Synthese, 190*(6), 997–1016. https://doi.org/10.1007/s11229-011-0055-x
Wilkenfeld, D. A. (2019). Understanding as compression. *Philosophical Studies, 176*(10), 2807–2831. https://doi.org/10.1007/s11098-018-1152-1
Yablo, S. (2020). Models and reality. In The Scientific (Ed.), *Imagination* (pp. 16–50). Oxford University Press.