How Idealizations Provide Understanding

Michael Strevens

To appear in *Explaining Understanding: New Essays in Epistemology and the Philosophy of Science*, edited by S. R. Grimm, C. Baumberger, and S. Ammon (eds.), Routledge, New York.

Abstract

How can a model that stops short of representing the whole truth about the causal production of a phenomenon help us to understand the phenomenon? I answer this question from the perspective of what I call the simple view of understanding, on which to understand a phenomenon is to grasp a correct explanation of the phenomenon. Idealizations, I have argued in previous work, flag factors that are causally relevant but explanatorily irrelevant to the phenomena to be explained. Though useful to the would-be understander, such flagging is only a first step. Are there any further and more advanced ways that idealized models aid understanding? Yes, I propose: the manipulation of idealized models can provide considerable insight into the reasons that some causal factors are difference-makers and others are not, which helps the understander to grasp the nature of explanatory connections and so to better grasp the explanation itself.

Can falsehoods produce understanding? The answer from science appears to be yes: many explanatory models contain idealizations, that is, deliberate falsifications, that are generally considered not to harm and even to help their users to understand why certain phenomena occur. Working with a particular view of the explanatory role of idealization and a particular view of understanding, this chapter attempts to explain how deliberate falsehoods play a part in helping scientists to better understand the workings of the world.

The thesis is then subjected to a criticism: although idealization may aid understanding, the objection runs, the same work can be done at least as effectively by an explanatory model containing nothing but the truth. The role of falsehood in understanding is acknowledged, then, but assigned an auxiliary status. Some arguments against this demotion are then formulated and endorsed.

1. Why Idealize?

Idealization is everywhere in scientific explanation. Accounting for the fact that a cannon ball and a musket ball dropped from the top of the Empire State Building hit the ground at the same time, for example, we might ignore air resistance—though it is certainly there. That is a fib, but a small one: the air resistance is slight.

Other idealizations involve a dramatic distortion of reality. To explain why dilute gases behave in accordance with Boyle's law, you reach for the ideal gas model, in which molecules are infinitely small and so never collide. That is hardly a negligible departure from reality. The same can be said of models in fluid dynamics that also assume infinitely small particles, or of models in population genetics that assume infinitely large populations (in which there is no genetic drift), or economists' use of models that represent economic agents—humans—as perfectly rational.

Do these fictions impede the explanatory power of the models? On the contrary, they enhance it, I argue in Strevens (2008), chap. 8. The role of idealization, according to the view propounded there, is to indicate that certain factors make no difference to the phenomenon to be explained—that collisions make no difference to gases' Boylean behavior, that genetic drift makes no difference to certain evolutionary events, that human irrationality makes no difference to certain economic phenomena.[1]

1. For some other approaches to idealization, see Weisberg (2007) and Strevens (2008), §8.11. Elgin's (2007) account of the role of idealization in understanding has some affinities to the story offered here, although Elgin does not emphasize a difference-making conception of explanation or the communicative role of deliberate falsification.

Behind this claim about idealization is an approach to explanation: the causal difference-making approach. In what follows I sketch only the basic form of the approach, as I will attempt to theorize about idealization without presupposing any particular implementation.

Consider an event that you might want to explain—the extinction of a species, perhaps. An explanation says something about the causal history of the event, but not everything that can be said. Rather, only aspects of the causal history that made a difference to whether or not the event occurred earn a place in an explanatory model. Factors that merely made a difference to *how* it occurred—that it happened quickly, that the organisms in this quadrant were the first to disappear, that the last of the line died on a Monday—are taken to be explanatorily quite irrelevant.[2] Or at least, they are irrelevant unless they also made a difference to the fact *that* the extinction occurred, in the sense that, were they somehow removed from the picture, the species would after all survive. (Different versions of the causal difference-making approach differ on how to understand removal and its consequences.)[3]

The same goes, more or less, for the explanation of a regularity such as gases' conforming to Boyle's law. There is a generic causal history to a gas's behaving in a Boylean manner, captured by a causal model of the gas that maps everything about the causal process by which, when held at a constant temperature, it follows the Boylean curve (along which, as the law dictates, pressure is inversely proportional to volume). An explanation, according to

2. Such factors are difference-makers for finer-grained states of affairs. If you want to explain why the extinction occurred quickly, then you must cite whatever additional aspects of the causal history made a difference to the fact that it occurred quickly, some of which factors presumably did not make a difference to the fact that it occurred at all—though of course, they made a difference to how it occurred.

3. The causal difference-making approach is taken, in various forms, by Lewis (1986), Salmon (1997), Woodward (2003), and Strevens (2008). Lewis and Woodward take a counterfactual approach to removal; Salmon consults a probability distribution; Strevens prescribes a process of progressive abstraction in a representation of the causal history, declaring a factor to be a difference-maker if, roughly, when the factor is abstracted away the description ceases to be a model of the causal production of the event to be explained.

the causal difference-making approach, describes just those aspects of the causal process that make a difference to the fact that a gas follows the curve. Declared to be irrelevant, then, are aspects of the gas that merely affect *how* it follows the curve—for example, how long it takes to relax after a change in pressure or volume.

As it happens, many salient properties of gases make a difference to how a gas follows the Boylean curve but do not make a difference to the fact that it follows the curve. The long-range attractive forces between molecules do not make a difference *that*. The transfers of energy between the gas particles and the container walls do not make a difference *that*. The very collisions between molecules do not make a difference *that*. A good explanatory model for Boylean behavior, then—a model that describes only properties of gases that are explanatorily relevant to Boylean behavior—will omit all of these things from its description of the gas, replacing them with an abstract specification of the gas's properties that is instantiated both by gases that have collisions and those that do not, by gases in which there are long-range attractive forces between molecules and those in which there are not, and so on.

Such a model must put certain limits on parameters such as molecular size and the strength of long-range forces, however: it must specify that neither is too large. It is not really the case, then, that nothing about the forces makes a difference to Boylean behavior; rather, the only thing about them that makes a difference is their not being too large—where one way to "not be too large" is to be zero, that is, not to exist at all.

In short, on the causal difference-making approach, an important part of explanation is to draw a clear line between difference-makers and non-difference-makers. Idealization accomplishes this task in a showy way: it falsifies certain non-difference-makers, by setting a variable that represents them to an extreme or default value.

Here, then, is how to interpret the ideal gas model, when it is proffered as an explanation of gases' Boylean behavior. Some of what the model says about gases is true. They are made of fast-moving particles that take up very little space. They expand to fill their containers uniformly. The pressure they exert on their container walls is the force exerted by particles colliding with the walls. These things are all difference-makers for, hence explainers of, Boylean behavior. Some of what the model says is false: that the particles are infinitely small and do not collide, that they exert no long-range forces on each other, that their collisions with container walls are perfectly elastic, and so involve no transfer of energy. These fictional claims—all involving extreme or default values—are telling you, in an oblique way, that certain properties of real gases do not make a difference to Boylean behavior. Molecular size (if not too large) is not a difference-maker. Hence, collisions (which require a non-zero molecular cross-section) are not difference-makers. Long-range forces are not difference-makers. Particle-wall energy transfers are not difference-makers.

Idealizations across the sciences should be interpreted in the same way. By falsifying a causal factor, an idealized model tells you that the factor is not a difference-maker for the phenomenon to be explained, thus helping you to draw the distinction between what is relevant and what is not. This helps you to better understand, I propose, why the phenomenon occurred or obtains. To make a clear case for this claim, however, I need to say something more about the nature of "understanding why".

2. Understanding Why

To understand why an event occurred, or why a regularity obtains, is to grasp a correct explanation of that event or regularity. This I call the simple view of understanding. The simple view is not a complete theory of understanding—it has nothing to say about forms of understanding that I call "understanding that" and "understanding with", or such other forms of understanding as may exist (Strevens 2013). But it is, I believe, a complete theory of "understanding why".

The term "explanation" can pick out at least two different, though related, things. Consider an explanation of Boyle's law. On the one hand, you can think of the explanation as a representation of the explanatorily relevant factors—that is, the properties of a gas in virtue of which it exhibits Boylean behavior. An explanation in this sense may appear in the pages of a textbook, will be written in a certain language, and does not exist until scientists come upon the scene and compose it. On the other hand, you might think of the explanation as what is represented: a set of properties and causal relations. In this sense, an explanation is not a representation; it is not in a book but out in the world. It exists before science; it is what scientists, when they arrive, undertake to discover.

It is a mistake to attempt to dictate which of these conceptions of explanation is correct, or even, I think, which is conceptually prior. We use the word 'explanation' in both ways, and accounts of explanation may usefully address either sense. There is a right and a wrong way, however, to interpret the word in my formulation of the simple view of understanding: the right way is the second way. Understanding why is a matter of grasping facts about the world out there—it is a matter of grasping, roughly, the causes of the phenomenon to be explained, and the facts in virtue of which they are causes.[4] To understand Boyle's law, for example, is to grasp certain properties of gas molecules and their causal consequences. To understand a species' extinction is to grasp certain aspects of the species' causal history, together with their causal connections. And so on.

What, then, is "grasping"? It is, clearly enough, something like knowledge. But it is stronger than knowledge.[5]

Here is a fact: trilobites are (or were) a kind of arthropod. To know this fact is, roughly, to be justified in believing it. (I take justified true belief as a

---

4. The simple view could, with a little extra work, be reformulated using the "representation" sense of explanation; the result would be somewhat more complex.

5. Whether it is strictly stronger, I will not say. Some writers have argued that in at least one respect, knowledge is stronger than what I here call grasp: grasp can be "lucky" in a way that knowledge cannot (Grimm 2006).

proxy for knowledge, then; the respects in which the two notions diverge are irrelevant to the point I wish to make.) A justified true belief that trilobites are arthropods is easily attainable even if you know very little about either trilobites or arthropods: it is enough to come across the fact in a trusted reference work. With justification you believe it; since it is in fact true, you thereby come to know it. But there is a sense in which, even though you know it, you may not understand it very well at all. You may have only a very hazy idea of what trilobites and arthropods are—respectively, say, some sort of long-extinct marine bottom-feeder and some invertebrate phylum in which legs are the rule.

I say in that case you do not fully grasp the fact that trilobites are arthropods. (You have a very partial grasp of the fact; grasping, then, comes in degrees.) Trilobite specialists, by contrast, certainly do grasp the fact (though perhaps not fully, if their information about trilobites and arthropods is in some crucial respects incomplete or mistaken).

What is the difference? Perhaps it can be characterized entirely in terms of knowledge: to have a fairly complete grasp of the fact that trilobites are arthropods requires a certain amount of knowledge about trilobites and arthropods. But what if that knowledge is itself incompletely grasped? Are there some very basic facts, knowledge of which is sufficient for grasp of which? Then perhaps grasping can be founded in knowledge of the relevant basic facts. Otherwise, it seems, it must be founded in some other epistemic or mental state.

Some of the literature on understanding suggests that grasp bottoms out in cognitive capacities (Grimm 2016 and for a related view, de Regt and Dieks 2005). My grasp of certain basic facts is constituted, on such a view, by my facility in making inferences about, or using, those facts. If my grasp of the trilobite fact bottoms out in grasp of the basic facts, then it too will be constituted, ultimately, by some kind of inferential capacity—the sort of thing, presumably, display of which would get me a job in paleobiology. Then again, you might think that this story gets the order of dependence precisely wrong: it is the capacities that are grounded in the epistemic state. It is my grasp of the basic facts about trilobites and arthropods, in other words, that gives me the ability to reason about them in sophisticated ways.

To give a philosophical account of grasping would be an extraordinary thing; I will not attempt it here. I hope that I have given you some sense of what I am talking about: to grasp a fact is like knowing the fact, but it involves a more intimate epistemic acquaintance with the state of affairs in question.

Before I move on, let me sketch an argument that grasping, not merely knowing, a correct explanation is necessary for understanding why. Consider the full general relativistic explanation for the earth's orbiting the sun. Among its elements: the mass of the sun; its impact on space-time curvature; the earth's tendency to trace a geodesic path in space-time; the fact that the geodesic is a corkscrew through space-time corresponding to the orbit we observe. I can know all these facts and I can know that they are related in the right ways to form an explanation of the orbit—that the curvature can be deduced in the right way from the local distribution of mass and the field equations of general relativity, and that the geodesic in such a curvature is the corkscrew—in the same shallow way that a trivia expert might know that trilobites are arthropods. I have the beliefs, and the beliefs are justified and true, but I comprehend only dimly the content of the beliefs. I can reproduce the explanation, but I do not really understand how it works—and so I do not really understand what it explains, the orbit of the earth. What is missing is a proper grasp of the explanatory states of affairs and their connections.

Understanding, then, thrives on "grip". Let me now put these ideas about understanding together with the previous section's ideas about the role of idealization in explanation.

3. The Role of Idealization in Understanding

3.1 Understanding: A To-Do List Idealizations help us to understand a phenomenon by giving us a better grasp of a correct explanation of the phenomenon—that is what I will now propose.

An explanation of a phenomenon, on the causal difference-making approach assumed in this chapter, is what you might call a "difference-making structure": a set of difference-makers for the phenomenon—properties of or events within a causal network in which the explanandum is embedded—together with the causal relations by way of which they make their difference, which I assume for the sake of the argument are captured or constituted by causal laws. To understand why a phenomenon obtains, then, is to get a grip on a certain difference-making structure, that is, to achieve the following sorts of goals: to grasp which aspects of the relevant causal process or history make a difference to the phenomenon; to grasp the way in which they make the difference that they do; to grasp the reason that other aspects of the relevant causal process or history do not make a difference to the phenomenon. Let me organize this to-do list for the prospective understander into two tasks:

1. Identify the difference-makers and (thus) the non-difference-makers.
2. Grasp the reasons why they are difference-makers and non-difference-makers.

The second task, I am supposing, is a matter of grasping the causal laws by which the difference-makers produce the phenomenon to be explained, and seeing that nothing further is required for its production.

How does idealization lend a hand? Most obviously, by clearly marking non-difference-makers—by planting, as it were, a big bright flag on top of causal factors that might have been expected to make a difference but do not, a flag saying "Despite what you may think, I am not a difference-maker!"

This is a useful, but rather rudimentary, contribution to understanding. It does not help the understander to deepen their grasp of the difference-makers, going from surface knowledge (as in the case of the biological ingenue's knowledge that trilobites are arthropods) to something more profound. It clearly distinguishes what needs to be deeply grasped, but that is all. Nor, apparently, does it help the understander to deepen their grasp of the causal laws by which the phenomenon is produced, and so to deepen their grasp of the reasons that the difference-makers make their difference. It is purely preparatory, a first step on the long road to understanding.

Or so you might think. In the remainder of this chapter, I take a more careful and more favorable look at the merits of idealization as a promoter of understanding.

3.2 Canonical Explanatory Models Must a good explanatory model contain idealizations? Evidently not: the role of an explanatory model is to represent a difference-making structure, and that can be done without perpetrating any falsehoods at all, by what I will call a canonical explanatory model.

A canonical model is simply a specification of a difference-making structure. It enumerates the difference-making aspects of the relevant causal process along with certain causal laws, and it derives the phenomenon to be explained from the difference-makers and the causal laws in a way that mirrors the process by which the difference-makers cause the phenomenon. A simple canonical model might explain the occurrence of an event by deriving it from certain initial conditions and a causal law; in so doing, it represents the process by which the initial conditions caused the event. More complex models will use some combination of laws or other regularities of the right sort; the details will vary to some degree with the specific account of explanation.[6] But they do not matter here.

The non-difference-making status of an aspect of a phenomenon's causal history is indicated, in a canonical explanatory model, by its absence. What is not in the model is not a difference-maker, with two qualifications.

6. My own way of fleshing out the story is presented in Strevens (2008), chap. 3.

First, a canonical model almost never gives a complete list of difference-makers; rather, it gives the difference-makers in a certain portion of the relevant causal process (Strevens 2008, §4.31), typically specified at a certain level of detail that abstracts away from fundamental-level processes (Strevens 2008, §5.4). These limits are tacitly agreed upon by both producers and consumers of canonical models. That a canonical model does not mention some event or some detail that lies outside or below the model's jurisdiction does not reflect either positively or negatively on the factor's difference-making status.

Second, some causal factors that are not explicitly described by a model might nevertheless have their difference-making status implicitly entailed by the model. Suppose—to take a toy example—event A causes event C which in turn causes the event to be explained E. A canonical model might cite A and two causal laws, one sufficient for A to cause C and another sufficient for C to cause E. It does not spell out the fact that C occurs, but it implies its occurrence as a part of the causal process leading to E. In so doing, it implies that C is a difference-maker for E.

Some further qualifications can be found in Strevens (2008), chap. 4; there is, however, no need to dwell on these matters.

What does a canonical model for the explanation of, say, gases' Boylean behavior look like? It includes all the true parts of the ideal gas model, specifying that gases are made of small, uniformly distributed molecules flying around a container and exerting pressure on its walls by way of their collisions. Where the idealized model ostentatiously departs from reality, however, the canonical model is more restrained. The ideal gas model sets molecular size to zero. The canonical model says rather that molecular size is very small. It does not specify any particular small size, because the exact size is not a difference-maker; rather, the difference-maker is the fact of the size's falling within a range from zero to some fairly small number. That molecular size does fall within that range (in the gases to which the model applies) is what the canonical model asserts.

The same goes for various other phenomena falsified by the ideal gas model: where the idealized model has a falsehood, the canonical model has a fact, but typically a "loose" fact: that the long-range attractive forces between molecules are, whatever their exact magnitude, very small; that the energy transfers between molecules and walls balance out, whatever their exact magnitude, so that there is no (or very little) net transfer. In short, where the idealized model flags some non-difference-makers with overt fictions, the canonical model carefully circumscribes all non-difference-makers by painstaking specifications of exactly what does make a difference.

3.3 Canonical Models Versus Idealized Models How do canonical models and idealized models differ in the way that they provide understanding? Let me tackle this question with reference to the to-do list presented above. The prospective understander of a phenomenon must, according to the list:

1. Identify the difference-makers and (thus) the non-difference-makers for the phenomenon.

2. Grasp the reasons why they are difference-makers and non-difference-makers.

I will proceed by constructing a case against idealized models: they are strictly less helpful in aiding understanding than are canonical models. I then consider some replies on behalf of the idealizer.

The first task is to distinguish difference-makers and non-difference-makers. Both models mark this distinction clearly in their own ways: fact versus fiction for idealized models; presence versus absence for canonical models. In drawing the line, they provide the prospective understander with the kind of preparatory help identified above: not help in grasping the nature of the difference-makers, but help in seeing which things are difference-makers and so need to have their nature grasped.

The second task has two parts: on the one hand, to understand why the difference-makers are difference-makers—to see how, together with the causal laws, they produce the phenomenon in question, and to see that their presence is necessary for this production—and on the other hand, to understand why the non-difference-makers are non-difference-makers, seeing that their presence is not needed to produce the phenomenon.

The canonical model provides considerable help in grasping non-difference-making status. It derives the phenomenon from difference-makers alone; in showing that the non-difference-makers are not needed for the derivation in the model's representation of reality it shows why they are not needed for the production in reality itself.

(Here I assume a derivation-based test for difference-making: a causal factor is a difference-maker for a phenomenon if it plays an essential role in the derivation of the phenomenon in some veridical causal model. The test is fleshed out and developed further in Strevens (2008), where it is asserted to be the ultimate criterion for difference-making. For the purposes of this chapter, however, it need be no more than a reliable heuristic, which I suppose even advocates of alternative accounts of difference-making will allow.)

The canonical model is less helpful for grasping difference-making status. It does show how the difference-makers play a role in producing the phenomenon in question, by deriving the phenomenon from the difference-makers. But such a derivation does not in itself show that any particular difference-maker is necessary for the production. For that, the prospective understander must do some additional work, asking themselves: What would happen if I removed this difference-maker from the story? Would the derivation still go through?

Now, the idealized model. It derives the phenomenon to be explained from a mix of real difference-makers and fictional non-difference-makers. Take the fictional non-difference-makers away (substituting veridical statements such as "Long-range forces are small") and you can still derive the phenomenon.

This shows that the non-difference-makers are non-difference-makers, but you must do the work yourself: in the same way that the canonical model does not help you to grasp the difference-making status of the difference-makers, the idealized model does not help you to grasp the non-difference-making status of the non-difference-makers. Worse, it does not help with the difference-makers, either. As with the canonical model, you are on your own, experimenting with subtracting difference-makers from the model in order to grasp the reasons that the derivation is not possible without them—and in so doing, to grasp the reasons that they are difference-makers.

In summary, the canonical model and the idealized model are on a par with respect to the first task, of distinguishing difference-makers and non-difference-makers, but the canonical model is strictly more helpful with respect to the second task, helping the prospective understander to grasp why the non-difference-makers have their status, if not why the difference-makers have theirs. Why, then, in pursuit of understanding, should we construct idealized models when we could construct canonical models that appear to be considerably more useful?

3.4 In Favor of Idealized Models Number one of three advantages that idealized models have over canonical models is simple, practical, and straightforward. The prospective understander's two tasks must be performed in the order given: first identify the difference-makers and non-difference-makers, then grasp why they are, respectively, difference-making and non-difference-making. Stumble in the first step and whatever effort goes into the second step is, at the very least, largely ineffectual. The first priority, then, is to get the first step right.

Arguably, the idealized model allows the prospective understander less scope for error than the canonical model. Its dramatic fictionalization of the non-difference-makers is harder to miss than the canonical model's quiet omission (Strevens 2008, 321). There is a cost to lighting up the non-difference-makers in this way—a less perspicuous presentation of the reasons for their non-difference-making—but it may be worth paying if the audience would otherwise lose the thread before the second act.

A second advantage of the idealized model over the canonical model is a relative logical or mathematical simplicity. To derive the Boylean behavior of a gas from the ideal gas model, you use very simple fictional values for various properties of the gas: zero molecular size, zero long-range forces, zero energy transfer in molecule-wall collisions, and so on. To derive Boylean behavior from the canonical model you must deal with certain ranges: molecular size between zero and some small number; long-range forces with magnitude between zero and some small number; and so on. This makes for a considerably more difficult derivation. Grasping how the difference-makers work together to produce the phenomenon to be explained, then, is easier with an idealized model.

The third advantage of idealized models builds on the second. Suppose, as I have been presuming, that the derivation test is used to determine the facts about difference-making. Then grasping causal factors' status as difference-makers or non-difference-makers is a matter of exploring various derivations or attempted derivations of the phenomenon to be explained, derivations that differ with respect to which causal factors are present and which are not. But which derivations to explore? The idealized model provides some salient suggestions as a well as a relatively tractable framework for trying them out.

As an explanation of Boylean behavior you are offered, let me suppose, the ideal gas model. The idealized model tells you what does and does not make a difference, but how can you use the model to better understand why? How to gain a better grasp of the way in which differences are made, or not made? Attempt some derivations of Boylean behavior. The ideal gas model supplies one set of values to try, mostly zeroes. Your knowledge of the properties of real gases supplies another set, which you can substitute into the ideal gas model, creating more realistic models—models in which molecular size is small but non-zero and collisions occur constantly, models in which long-range forces are small but non-zero, and so on. (You can even fabricate values for force strength and other idealized quantities, since you know that the exact values do not matter.) See which factors can be removed without affecting the derivation of Boylean behavior and which cannot. The model already supplies the answers, of course—it distinguishes difference-makers and non-difference-makers—but in working out these answers for yourself, you also see *why* some factors are necessary for the derivation and some are not. There lies your understanding.

There is a caveat: that a factor makes no difference to a derivation when other factors take especially simple forms—in particular the "zero" form that represents their absence—does not guarantee that they are not difference-makers. That small long-range forces make no difference to Boylean behavior in a gas with no intermolecular collisions, for example, does not entail that they make no difference in a gas with collisions (though as a matter of fact, they do not). Full understanding of difference-making thus requires, in principle, that realistic values for non-difference-makers be tried out in combination as well as separately. In practice, however, the reasons for non-difference-making are often the same in isolation as in combination. Molecule-wall energy transfers make no difference because in the long run they cancel out; this is true regardless of collisions and long-range forces. There is plenty of understanding to be had, then, testing just the simplest variations of the idealized model.

Could you do all of this with the canonical model? Yes—you can achieve any level of understanding by contemplating a canonical model alone—but it is a far more complex operation. Even sophisticated scientific practitioners more easily grasp the reasons things do and do not make a difference to a phenomenon by manipulating an idealized model. And that is how idealization aids the pursuit of understanding.

References

de Regt, H. W. and D. Dieks. (2005). A contextual approach to scientific understanding. *Synthese* 144:137–170.

Elgin, C. (2007). Understanding and the facts. *Philosophical Studies* 132:33–42.

Grimm, S. R. (2006). Is understanding a species of knowledge? *British Journal for the Philosophy of Science* 57:515–535.

———. (2016). Understanding and transparency. In S. R. Grimm, C. Baumberger, and S. Ammon (eds.), *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*. Routledge, New York.

Lewis, D. (1986). Causal explanation. In *Philosophical Papers*, volume 2, pp. 214–240. Oxford University Press, Oxford.

Salmon, W. C. (1997). Causality and explanation: A reply to two critiques. *Philosophy of Science* 64:461–477.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press, Cambridge, MA.

———. (2013). No understanding without explanation. *Studies in History and Philosophy of Science* 44:510–515.

Weisberg, M. (2007). Three kinds of idealization. *Journal of Philosophy* 104:639–659.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, Oxford.