Philosophical Review

Collin Rice, *Leveraging Distortions: Explanation, Idealization, and Universality in Science*. The MIT Press, 2021. Pp. i-xii, 1-353.

Reviewed by H.K. Andersen, Simon Fraser University

Questions about idealizations in science are often framed along the lines of, How can science be so effective when it gets so much wrong? Rice's book, *Leveraging Distortions: Explanation, Idealization, and Universality in Science* offers a refinement on this framing, where we need not commit to the premise that idealizations are, in fact, wrong, that they need to be contained to the irrelevant parts of a model, or should be explained away as mere appearance. Rice takes a holist approach in which idealization is more like a process by which models as a whole are leveraged into better fit with their targets. Idealizations should not be carved out one by one on this approach; they make sense in the context of the models in which they figure, and they distort in ways that illuminate features like universal behavior in the systems being modeled. This is a refreshing approach to how idealizations work, one which does not require the common presupposition that idealizations are simply false.

By *universality*, Rice mean " the stability of certain patterns or behaviors across systems that are heterogeneous in their features. Universality classes are, then, just the group of systems that will display those universal patterns or behaviors." (155). Universality enables a more abstract description of systems than what scientists may have started with, and this process of making the description of the behavior more universal serves to identify common causal structures implemented in very different physical mediums. Different descriptions of causal relata facilitate identification of more unifying patterns of behavior. Given how often philosophers think of abstraction as somehow eliminating causation, by identifying causation too strongly with microphysical details, universality is a helpful way to bring the process of abstracting description back into contact with the way in which models inevitably involve causal structure, and how that causal structure itself can be better understood by connecting classes of systems with heterogenous physical media and similar behavior, by showing how the more abstract descriptions of causal structure are deployed in each.

There are two specific features of his view that set Rice's book apart from most other contemporary views on idealizations. The first is the explicit emphasis on holism. Often, idealizations are isolated from models and then assessed on their own, after extraction from the modelling context in which they were made. In evaluating idealizations as individual propositions removed from surrounding context, it is somewhat unsurprising that many look inaccurate. Rice aptly shows how idealization plays a key role in identifying universality behavior by distorting a whole, undecomposed, model. This focus on holism and the role idealizations plays in a larger modeling context helps Rice's treatment of idealizations stands apart from many others, including those he explicitly engages with such as Potochnik (2017), Strevens (2011), and Khalifa (2017). This approach fits better with the usage of idealizations in science by not needing to explain *away* the widespread reliance on idealization in so many

sciences. Even if one thinks the other accounts are successful in trying to explain why idealizations can be used in science despite falsity and misrepresentation, there is something uncomfortable about explaining such widespread use of them by framing it as apparently irrational. Rice's account does not require starting from a framing where scientists rampantly engage in apparently irrational practices, and then explain why it is not as bad as it looks. Instead of using idealizations *despite* falsity, idealizations are part of a coherent package that can be used for explanatory leverage.

The second feature that sets his view apart follows from this: idealizations are a tool to be actively used, not peculiarities to be explained away or dubious commitments to be minimized. Too often, idealizations are treated as some kind of representational failure, a compensation for epistemic limitations. In a more epistemically perfect world, on such thinking, idealizations could be done away with. Rice turns this around: idealizations are not something we put up with or have to be resigned to; they are a key tool to be used in positive ways to generate explanations and for building bodies of understanding. This is where the 'leveraging' part of the title comes in: idealizations are actively relied on to achieve modeling techniques that would be impossible otherwise. They are a lever by which to torque a model into better alignment. This positive feature of idealizations accounts for the advantageous character of idealizations as a feature, not a bug.

While Rice is, in my view, exactly right to reject these background presuppositions about the falsity of idealizations, I would also add that he could go further in this regard; the book would benefit from more explicit discussion of what he means by truth or falsity. There are pragmatist versions of truth, for example, that are quite consonant with his final view, so that it need not be framed as a puzzle that *false* statements somehow work to return genuine knowledge. Idealizations are usually presupposed to be false; authors like Potochnik (2017) in fact define them as false, such that if it is an idealization, then by definition, it could not be true. Rice does not seem to endorse this, yet accurate representation is left hanging somewhat. A discussion of epistemic standards of veridicality that should be used for the holistic evaluation the of models, and the ways in which various identifiable components of those models accomplish this without decomposition would strengthen his overall push towards a more explicit and foregrounded holism about models and his claims in chapter 8 about realism.

That is quite mild, as critical remarks go, and most of the book is full of detailed examples and other discussions that don't require a further discussion of truth. There is a lot covered in this book, much of which Rice has written about elsewhere, and some of which he extends, refines, or adds to in new ways in the book. In the introduction, Rice stakes the main claim that pervasive distortion doesn't just *happen* in science, it is *central* to science working as well as it does that such distortion take place. This sets up the later chapters on universality as a behavior that can be instantiated in physically heterogenous systems and identified with more abstract (and distorting) descriptions of those systems. This introduction does a good job of situating why this alternative stance towards idealizations as pervasive distortions that are used for purposes that cannot be served with other tools differs from approaches where idealizations

are considered after isolating them from modelling contexts and then evaluating them as false yet useful.

Chapter 2 discusses what Rice calls the causal or causal-mechanical paradigm in literature on explanation. The causal approach, as he characterizes it, explains an event by giving the relevant factors in the event's causal history. Salmon, Woodward, Strevens, Potochnik, and new mechanisms are highlighted as examples of this. Rice is right to highlight how widespread discussions of causation are in discussions of explanation, and it is great to see Salmon given more credit. At the same time, this chapter lumps together some heterogeneous approaches, like Woodward's (2005) account of causal explanation, for example. Woodward gives an account of those explanations that are causal, without claiming that this is exhaustive of all explanation; there could be non-causal explanations, he just isn't discussing it. Strevens (2011), in contrast, takes himself to be providing a complete account of explanation based on causation; Potochnik (2017) as well offers an account of explanation in which causation, in the form of causal patterns, play a necessary role.

Chapter 3 follows this up by demonstrating with a series of examples a number of explanations that do not involve causation. This chapter may be overkill if the goal was to demonstrate that not all explanations need be *causal* explanations, since some of the apparent targets, like Woodward, already agree with this, and there is a lot of interesting work on distinctively mathematical explanations that highlights how they contrast with and complement causal explanations that he does not engage with. But as a collection of examples of non-causal explanation, this chapter has new material to add to existing examples, especially to the examples of distinctively statistical explanations given by Lange (2016).

In chapter 4, Rice lays out his own Counterfactual Account of explanation, and contrasts it with other such accounts. He offers three criteria that any such account should meet that will be useful in these discussions (93), even if one does not want to adopt Rice's own particular account. The details of Rice's own account here seem compressed, and if one just reads this chapter, it is hard to see how this is supposed to work and be a genuine move forward. The later chapters, especially 6, 7, and 9, show how the account works when applied, which is illuminating. It would thus be useful, for instance if teaching from the book in a seminar, to pair chapter 4 with one of these further chapters, especially Chapter 6.

Chapter 5 is brief, focused on how decomposition of models into sub-components that are then treated separately simply doesn't work for most models. Rice makes some very clear points about why models must be treated holistically, solidifying his point about idealizations as distortions in those models which don't make sense when taken out of that context through attempts at decomposition.

Universality, a term of art here that follows on Rice's other work (inter alia, Rice 2018 and 2019, Batterman and Rice 2014), is given detailed treatment in Chapter 6. This chapter lays out some detailed case studies, and illustrates how the holistic distortion involved in idealization is what conveys or captures the specifically modal information in a model. Chapter 7 continues with

themes Rice has written about elsewhere, multiscale models and how universality fits into considerations of scale and renormalization.

Chapter 8 moves on to consider how models can provide understanding even when they do not do so by providing explanations. Rice's examples involve cases where scientists have incomplete explanations, so some might consider these to be explanations already, since one need not require that an explanation be fully complete in order to count as an explanation. This chapter also connects understanding to realism and scientific progress. Idealizations have often been treated as failures for realism, where an otherwise successful model is purportedly decomposed into elements, some of which are clearly not literally representationally accurate, in the way one might suppose necessary to be a realist about that component (another way in which naive correspondence treatments of truth sneak into philosophy of science, by way of assuming that bits of models should map one to one to bits of the world, and that realism about a model fails if there are idealizations that don't map in this simplified way). He draws on his own account of Factive Understanding, in the first part of the chapter, to lay out an alternative approach to realism where the focus is not on isolated model components but on the body of understanding that models produce for scientists. This body of understanding, which again requires holism, can serve as an epistemic basis for realism about the behavior thus understood.

Finally, in chapter 9, Rice brings together all the themes in the book and makes the clearest case yet for how idealizations are used as 'holistic distortions' that are not merely part of science, but central and positively contributory to the success of modeling techniques in providing both explanation and understanding. This chapter is a great conclusion to bring together the different topics in the book. Many of the other topics are ones Rice has written about elsewhere, and this concluding chapter helps make sense of the synoptic project into which all this work fits. If one were teaching with this this text, this might be a good chapter to start with, rather than end with.

Overall, this book does a nice job of bringing together Rice's previous work while also extending that work with new examples and ones worked out in more detail, and of connecting the different topics in a cohesive way around the orientation towards holism and idealizations as holistic model distortions. This makes it a great addition to a range of contemporary discussions around explanation, models, understanding, realism, and a good starting point for graduate students to get into these topics.

References:

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

Khalifa, K. (2017). *Understanding, explanation, and scientific knowledge*. Cambridge University Press.

Lange, M. (2016). *Because Without Cause: Non-Casual Explanations In Science and Mathematics*. Oxford University Press.

Potochnik, A. (2017). *Idealization and the Aims of Science*. University of Chicago Press.

Rice, C. (2018). Idealized models, holistic distortions, and universality. *Synthese*, *195*(6), 2795-2819.

Rice, C. (2019). Universality and the problem of inconsistent models. In *Understanding Perspectivism* (pp. 85-108). Routledge.

Strevens, M. (2011). *Depth: An account of scientific explanation*. Harvard University Press.

Woodward, J. (2005). *Making things happen: A theory of causal explanation*. Oxford university press.