



## Universality is not universal: how much can we explain with falsehoods?

Collin Rice: *Leveraging distortions: explanation, idealization, and universality in science*. Cambridge MA: The MIT Press, 2021, 353 pp, \$65.00 PB

Martin King<sup>1</sup>

Accepted: 10 March 2022  
© The Author(s) 2022

In *Leveraging Distortions*, Collin Rice brings together a lot of previous work in arguing that the practice of scientific modelling is rife with deliberate misrepresentations and that, contrary to extant accounts of explanation, this is integral to how science is able to generate understanding. The aim of the book is to make philosophers confront the positive role of idealisation in formulating explanations. Rice pits himself against almost all views of explanation and modelling in that he gives up on the notion that truth and accurate representation are aims of scientific modelling and are required for good explanations.

Many accounts of explanation take, as a starting point, the fact that science necessarily idealises. For many, this raises the question as to how models with falsehoods and inaccuracies can provide genuine explanations and generate factive understanding. This is often answered, according to Rice, by making assumptions about the separability of different aspects of models—that the idealised features of models can be ignored while we can focus on the accurate or approximate features that, in any case, do the explanatory work. Rice's claim is that the idealisations, or distortions, are holistic and pervasive and often the very features central to the explanation are the ones being distorted. Not only is this not a problem for Rice, but it is actually the reason why we are able to use various mathematical modelling techniques that expose the high-level counterfactual dependencies that scientists care about.

The book begins by reviewing popular causal accounts of explanation and stressing that a more general account of explanation should not restrict the domain of explanatory models to those that are accurate representations of causal dependencies. As many others have argued, there are a great deal of models that are considered explanatory by scientists but do not draw specifically on causes for their

---

✉ Martin King  
mking@uni-bonn.de

<sup>1</sup> University of Bonn, Bonn, Germany

explanations. In Chapter 4, Rice presents his counterfactual account of explanation that, while sharing certain features with Woodward's interventionist account, does not make any demands on the causal nature of the counterfactuals. He lays out three requirements that an account of explanation should satisfy, viz. it should capture causal and non-causal explanations, clarify the link between explanation and understanding, and account for the role of idealisation in explanatory models.

These requirements are intentionally not about the features a model should exhibit, in order to be explanatory, but features of a satisfactory account of explanation. This is an interesting shift in focus but will leave some readers wondering whether some further conditions on explanatory models should be introduced, given the broad scope of inaccurate and highly idealised models he is intending to capture. For Woodward, explanations need to track causal dependencies in order to prevent backtracking, i.e. to keep mere correlations that satisfy counterfactuals from being considered explanatory. Without this causal support, one wants to know how such counterfactuals are debarred.

To address this, Rice only says that "one just has to be careful to avoid [them]" (96).

In Chapter 5, Rice argues that models cannot be decomposed into relevant and irrelevant parts, such that one can preserve veridical aspects and ignore the idealised parts. These models are *holistically distorted*. In Chapter 6, he finally reviews some cases studies of models that use universality classes in order to focus on high-level dependencies. The use of universality also comes into play in Chapter 7, where he argues that it can help avoid conceptual issues related to models at different scales. The apparent problems of inconsistent models only arise because one takes the models to be accurate and truthful descriptions of target systems, but when one sees them as intentionally distorted models that aim only to reveal certain large-scale behaviours and counterfactual dependencies, the tensions dissolve. This raises some worries for his account that have only been dealt with in a promissory manner, until these later chapters.

In Chapters 8 and 9, Rice brings together many of the previous results and addresses questions about why one should think of these inaccurate models as explanatory. The answer, in brief, is that they provide understanding—an answer reminiscent of that given by Bokulich (2008) in defending the explanatory nature of highly idealised models. One does not get knowledge of true propositions with detailed and accurate representations of the difference-making features of models, such as one finds in Strevens (2008), for example. Rather these holistically distorted models give understanding by allowing for the mathematical treatment that exposes high-level counterfactual dependencies that would otherwise remain obscured. Rice calls the understanding "factive", as long as "most of what is believed about the phenomenon [is true]" (251), even though for many this would indicate a non-factive view of understanding. For Rice, it seems important to call this factive in order to stress that the beliefs about these counterfactuals are true, even if they come from models with falsehoods and deliberate misrepresentations. The idea that factive understanding can come from false models is an interesting and important contribution of the book.

Another success of the book is that the view of idealisation, understanding, and explanation are put to further use in articulating understanding-based notions of progress and of realism. There is no problem on this account for the proliferation of conflicting models, either synchronically in terms of models at different scales, or diachronically in terms of the advancement and progress of science. That the details of a given model conflict with another model or are eventually found to be false is no impediment to it being explanatory or providing factive understanding. Although, it must be said that this “understanding realism”, as he calls it, may not be of much comfort to someone who wants to be a traditional scientific realist. It even undermines the idea that features of good models should be interpreted literally and realistically. However, this “understanding realism” demonstrates that this view of idealisation and explanation hangs well together in a coherent picture of science in general.

The book indeed has many virtues, including highlighting the important role of idealisation and distortion for the formulation of some explanations and stressing the positive role of inaccuracies. However, one may worry that, for an account that embraces false models, it is not very concerned with avoiding backtracking counterfactuals, or more generally, with debarring non-explanatory models. The book project follows a recent tradition in the explanation literature that aims to be more and more inclusive of what counts as explanatory. While paying attention to scientific practice, and taking seriously what is considered to be explanatory by scientists is an important consideration, a philosophical account of explanation should not shy away from being somewhat normative and maintaining a distinction between prediction and explanation. One can expose large-scale behaviours by explicitly distorting features of a model, but that alone does not make the model explanatory. Along this line, the case studies that ultimately make the argument for this account lack the detail and depth to be thoroughly convincing, and the descriptions of other accounts of explanation will not serve as an introduction for someone who does not already know them. More importantly, however, one worries that *Leveraging Distortions* has leveraged some distortions of its own in portraying the practice of generating explanations as being largely about universality. The account is intended to be general but is designed around very specific explanations of multiple realisability. Batterman (2002) and Batterman and Rice (2014), for example, raised an interesting point: answers to questions about how a model explains a given phenomenon are not also answers to questions about how a class of models with different microconstituents can all exhibit the same behaviour. This point has been well acknowledged in the literature, but one must also keep in mind that the reverse is also true: explanations of universality are not at the same time explanations of individual phenomena. Universality explanations with minimal models are only one specific kind of explanation, and while these explanations may be found across a variety of scientific disciplines, this does not indicate that the account is actually very general. Universality is an interesting phenomenon, but sometimes we will want a detailed and truthful account to explain particular phenomena.

Ultimately, the book presents a contemporary view of explanation and understanding that attempts to shift the discussion in the philosophy of science away from

some false and simplistic assumptions about the role of idealisation in explanation. In this respect, I think the book is, and will be, quite successful.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Batterman, R.W. 2002. *The devil in the details*. Oxford: Oxford University Press.
- Batterman, R.W., and C.C. Rice. 2014. Minimal model explanations. *Philos Sci* 81 (3): 349–376.
- Bokulich, A. 2008. *Reexamining the quantum-classical relation: beyond reductionism and pluralism*. New York: Cambridge University Press.
- Strevens, M. 2008. *Depth: an account of scientific explanation*. Harvard, MA: Harvard University Press.
- Woodward, J. 2003. *Making things happen: a theory of causal explanation*. Oxford: Oxford University Press.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.