**Leveraging Distortions: Explanation, Idealization, and Universality in Science, by Collin Rice**

**Universality, Understanding, and Realism**


## Abstract

*Leveraging Distortions* makes a strong case for the epistemic importance of pervasively distorted models in science. I identify three areas where the argument falls short: the reliance on counterfactuals to fix dependence relations, the deployment of universality classes to vindicate the use of distorted models, the commitment to factivity of understanding. I suggest remedies to the problems I identify.



## Keywords

dependence relations, counterfactuals, universality classes, factivity


*Leveraging Distortions* makes a compelling case for the epistemic importance of pervasively distorted models in science, arguing that such models advance understanding not despite, but because of, their distortions. It maintains that scientific explanations need not be causal or mechanistic. Science aims to explain a variety of dependence relations, including mathematical, topological, and perhaps supervenience relations among phenomena. The recognition that not all dependence relations are causal is gaining considerable traction in metaphysics. Rice brings it to bear on the philosophy of science. This is a valuable contribution. In what follows I focus on three strands in Rice's argument. One concerns his confidence in counterfactuals; the second, his deployment of universality classes, the third, his argument for factivity. I identify what I take to be weaknesses in these strands and suggest ways to alleviate them.

Rice contends that the multiple dependence relations science relies on can all be captured by counterfactual analyses linking explanans with explanandum. Remarkably, he ignores the vast literature on counterfactuals and dependence relations (see Tahko and Lowe, 2020). This leads to trouble. His discussion of the periodical cicada provides an example of the difficulty. Periodical cicadas burrow into the ground immediately after birth. They emerge in enormous broods every 13 or 17 years. Their emergence cycles rarely coincide with the population cycles of their predators. That is an evolutionary advantage, since even if every predator eats its fill, plenty of cicadas remain to reproduce. The fact that 13 and 17 are prime numbers explains the advantage. The explanation admits of a straightforward counterfactual gloss. If the cicada's periodicity had been a composite number of years, their emergence would more frequently coincide with the presence of a maximal number of predators. That would make the species more vulnerable. Rice, however, favors a different explanation. He says, “I think we can agree that the counterfactual statement, 'If 13 and 17 were not prime, then the cicada's life cycles would have been different' is true” (p. 115). I do not agree, for I consider the antecedent inconceivable. What possibility are we supposed to entertain: that 17 was divisible by 2, by 3, by 5, by 7? If it were divisible by any of these, it would not be 17. Rather, the numeral '17' would denote a different number. It is inconceivable that 17 could fail to be a prime number.

Some rather drastic moves are available to shore up Rice's claim. Like Yablo (1990), he might maintain that inconceivability does not entail impossibility. Then he could argue either that '17 is a prime number' is contingent (see Hodges 2013) or that '17 is not a prime number' is true in impossible worlds (see Priest 1997, Rescher and Brandom 1980). Neither alternative is promising. If '17 is a prime number' is contingent, there are possible worlds where it is false. These possible worlds are apt to be so distant that they shed no light on viable evolutionary strategies in the actual world. The same holds a fortiori if to make sense of the claim we need to resort to impossible worlds.

Why, one might wonder, didn't Rice opt for my unobjectionable counterfactual? Evidently he thinks that the explanation must depend on the fact that 13 and 17 are prime. It is not enough that the periodicity and the evolutionary advantage are counterfactually correlated. The advantage, he thinks, rests on the fact that 13 and 17 are prime. Since counterfactual reasoning is reasoning about what if things had been different, he evidently thinks we can give a suitable counterfactual analysis only if we can intelligibly ask 'What if 13 and 17 had not been prime'? The problem is that counterfactuals are ill suited to explaining mathematical dependencies. For those dependencies are necessary.

Even when mathematical cases are set aside, the contention that counterfactual explanations can replace causal ones faces a familiar objection. The length of its shadow counterfactually depends on the height of the flagpole. If the flagpole had been taller, the shadow would have been longer. The converse holds too. If the shadow had been longer, the flagpole would have been taller. Still, we insist, the length of the shadow depends on the height of the flagpole, not conversely. An adequate theory of explanation should explain the asymmetry. Rice agrees. But he maintains that he can get the asymmetry by avoiding 'backtracking' counterfactuals. 'Changes in the length of the shadow (due to other factors, such as the sun moving) do not change the height of the flagpole.' (p. 96). This is so, but changes in the height of the pole would likewise leave the shadow the same length. Such changes are not unthinkable. A set designer might bring it about that the shadow's length remains fixed across changes in illumination during a scene by continually adjusting the height of the (off stage) flagpole. Rather than relying on counterfactuals to explain the dependence of the shadow on the flagpole, we apparently need to fix the dependence relation in order to decide which counterfactuals to credit.

Rice maintains that if a seemingly important feature of the explanandum is irrelevant, that it is irrelevant should be part of the explanation. This makes explanation contextual, since context determines which irrelevances must be included. Rather than saying that the height of the flagpole explains the length of the shadow, we need to say, the height of the flagpole explains the length of the shadow to audience *A* where their interests are *B*, *C*, and *D*. Science does not explain things simpliciter. It explains things to certain audiences with certain backgrounds and certain interests.

This insight underwrites Rice's use of universality classes. Universality classes are classes of real, possible, and/or model items that manifest the same pattern or behavior. They figure in explanations where the same macro-level behavior is displayed by phenomena with heterogeneous microstructures. Then micro-structures are irrelevant. Because such irrelevance is surprising, explanations should flag it. When causally and mechanically heterogeneous phenomena belong to a single universality class, Rice maintains, scientists can explain via a highly idealized model that also belongs to that class. Since the causal/mechanical details don't matter, the behavior displayed by the model can be attributed to, and used in an explanation of, the members of the class. Such a model can be drastically distorted, since its aptness depends only on the factors that it shares with other members of the class.

The discussion of universality classes is one of the most important contributions of *Leveraging Distortions*. Very roughly, an account that appeals to universality classes maintains the following: Anything that does *A* does *B*; items with heterogeneous causal and mechanical structures do *A*. So the explanation of their doing *B* does not depend on underlying causal or mechanical structures. Scientists can therefore model the relevant behavior via an idealization that prescinds from mechanical and causal factors. A model may drastically and pervasively distort properties by, e.g., representing a finite population as infinite; a huge planet as a point mass; etc. No matter. These distortions function to make calculations tractable without falsification. Since the factors they prescind from, and the distortions they make are irrelevant, it does not harm to ignore them.

Still, there is a worry. Every item belongs to a vast number of universality classes. Nor does membership in a particular universality class always require explanation. What determines which universality classes are worthy of attention? What determines why an item's membership in a given universality class matters? What determines which factors can legitimately be ignored? The real items in a universality class inevitably share irrelevant features. We need a basis for ignoring them. Prior to the use of a model, then, we have to grasp the importance of some of the behavior of the real items in the universality class and the factors that bear on that behavior. But then the model seems to presuppose rather than explain. There seems to be no way principled way to mark out the cases where a model's being a member of a given universality class enables it to explain the behavior of the heterogeneous real phenomena that are also members of that class.

To resolve this difficulty, Rice could invoke a variant on Strevens' account (2008). Strevens maintains that scientific explanations must be causal. Hence all difference-makers must make a causal difference. Rice, as we have seen, disagrees. Still, he might introduce the idea of a non-causal difference-maker. Then he could say that a drastically and pervasively distorted model can explain when its divergences from and distortions of the target are not difference-makers. Since it makes no difference that, e.g., Jupiter is not a point mass, a model that treats it as a point mass can explain its gravitational effects.

If he acknowledged the utility of exemplification, Rice could explain why seemingly significant features are non-difference makers. The model, through pervasive distortions, exemplifies explanatory features and relations that it shares with the target. It shows that the seemingly significant features that are ignored are in fact not difference-makers. Rice rejects this move because he thinks that exemplification 'struggles to accommodate the existence of multiple models that make contradictory assumptions about the same patterns or features of real systems' (p. 289). Not so. Different models, via different pervasive distortions, exemplify different features of the systems. One model represents the nucleus as a rigid shell; another as a liquid drop. A shell model exemplifies features of the nucleus that figure in the explanation of the stability of nuclides. A liquid drop model exemplifies features that figure in the explanation of binding energy. The selectivity of exemplification makes it suitable for explaining why the features that the liquid drop model are appropriately absent from the rigid shell model (see Elgin 2017). Different – indeed incompatible – distortions figure in the exemplification of the two sets of features.

Rice contends that despite the ubiquity of pervasively distorted models in science, scientific understanding is factive, since "in order to genuinely understand a natural phenomenon, *most* of what is believed about the phenomenon – especially about certain contextually salient propositions – must be true" (p. 251). That most of what is *believed* about the phenomenon is true may be correct. To believe that *p* is to believe that *p* is true. It may be plausible that we should only be realists about what we believe to be true. But if we extract truths from the pervasively false models that encapsulate our grasp of the phenomena, we leave their justification behind. We may believe truths, but we lack justification for them. True beliefs that lack justification do not constitute understanding. This suggests that understanding is a matter of acceptance rather than belief, and that it is embodied in the idealized models that science reflectively endorses (see Elgin 2017). Acceptance better captures the importance of models and the universality classes that Rice champions.

Despite these worries, *Leveraging Distortions* effectively argues that distortions are a source of epistemic power. Rather than reluctantly settling for distortions, science exploits them to devise models that advance understanding.

References


Elgin, Catherine (2017). *True Enough*. Cambridge MA: MIT Press.

Hodge, Wilfred (2013). 'Modality in Mathematics' *Logique et Analyse* Nouvelle Série, 56:5-23,

Priest, Graham (1997). 'Sylvan's Box: A Short Story and Ten Morals' *Notre Dame Journal of Formal Logic* 38:573-581.

Rescher, Nicholas and Robert Brandom (1980). *The Logic of Inconsistency: A Study in Non-Standard Possible World Semantics.* Oxford: Basil Blackwell.

Rice, Colin (2021). *Leveraging Distortions: Explanation, Idealization, and Universality in Science.* Cambridge MA: MIT Press.

Tahko, Tuomas E. and E. Jonathan Lowe (2020). 'Ontological Dependence' *Stanford Encyclopedia of Philosophy* Edward N. Zalta (ed.) <https://plato.stanford.edu/archives/fall2020/entries/dependence-ontological/>.

Yablo, Stephen (1990). 'Is Conceivability a Guide to Possibility?' *Thoughts: Philosophical Papers volume 1.* Oxford: Oxford University Press.