

Review of Angela Potochnik's *Idealization and the Aims of Science*

Daniel C. Burnston

Tulane University, Philosophy Department

Tulane Brain Institute

dburnsto@tulane.edu¹

Angela Potochnik's *Idealization and the Aims of Science* is a fantastic book. In it, Potochnik argues for a compelling, global picture of how science works – one that seeks to clarify how the practice of science relates both to human cognitive capacities and to the world we seek to understand. The book is ecumenical yet concise. It is broad but focuses on the details. It seeks to make generalizations about science, but it does so through diverse analyses of particular scientific practices. If one wanted a single book that summed up both the challenges and opportunities in current philosophy of science, one could hardly do better.

There are two concepts at the heart of the book: *idealization* and *causal patterns*. Idealization involves representing something falsely – as having properties it does not in fact have – for a purpose, for instance to explain or predict what the represented thing will do. Causal patterns are dependencies between factors, which are revealed under manipulation and “embodied” by phenomena we want to explain

According to Potochnik, idealization is inevitable. Science, she rightfully insists, is a human activity, pursued by real people with cognitive limitations. Scientists are further influenced by their explanatory goals, their values, and their conceptual and social backgrounds. The world, on the other hand, is a “multifactorial miasma” (74). Any phenomenon we're interested in is going to embody a huge number of causal patterns. It is going to be embedded in a complex nexus of influences, at many spatial and temporal scales, and with no clear boundaries between what affects it and what doesn't. Given the gap between the complexity of the world and humans' limited understanding, we should not expect, or seek, complete theories of phenomena. At best, we can hope to understand “focal” causal patterns, which we isolate for analysis by extracting them from the complexity of their surroundings.

Idealization, for Potochnik, is what allows for this isolation, and thus what fills the gap. Different scientific projects will focus on different causal patterns, representing different aspects of the phenomenon using different simplifying assumptions. Idealizations are thus both *rampant* and *unchecked*. They are rampant because *every* scientific field idealizes, and because *every* pattern is represented in an idealized way. They are unchecked because their use is not regimented by strict standards of justification. Moreover, since idealization is inherent to scientific representations and explanations, we should not hold out hope for “de-idealization.” Far from being a negative, Potochnik thinks that the success of science is largely predicated on idealization. It is what allows us to really access the world despite our limitations.

The first few chapters of the book lay out the general approach outlined above. Chapters four and five apply the framework to develop an account of explanation. According to Potochnik, scientific explanation does not seek the truth. Instead, it is an exercise in developing

¹ Thanks to Alican Başdemir, Andrew Buskell, Angela Potochnik, Lachlan Walmsley, and the editors at *Philosophy of Science* for extremely helpful discussion and comments.

understanding. We understand, and explain, a phenomenon when we successfully represent the causal patterns it embodies. This explanation is *adequate* when it “accounts for” the phenomenon. Successful accounting-for comprises two conditions: the posits in the representation must be “acceptable,” i.e., accurate enough for the aims of the researcher; and the representation must show that the phenomenon is to be expected under certain conditions. Accounting-for thus involves a kind of entailment of the phenomenon, but as Potochnik clarifies, this is a weak notion of entailment, on which the phenomenon is to be expected to whatever extent the simplifying assumptions made in the idealization hold.

Since idealization involves ignoring or simplifying potentially relevant causal factors, we should expect our explanations to be fraught with exceptions and limitations. We should also expect limited interaction between different scientific projects. Since representations are defined by their idealizations, and the focal patterns that these help to isolate, we can expect different fields and modelling projects to remain relatively isolated. “Integrations” – representations of causal patterns that draw on multiple models at once – should be rare, and are in many cases simply undesirable. Instead, what we have is a patchwork “division-of-labor” between different scientific projects focused on different patterns, even when those projects seek to explain the same phenomenon.

The last few chapters extend the framework to give a global picture of the organization of scientific endeavor. Scientific “products” – e.g., explanations – do not fit into any clear order or structure, according to Potochnik. Scientific fields do not focus strictly on particular levels, and there is no universal notion of levels that will provide systematic ordering to their products. While we may be able to, in certain circumstances, define levels according to compositional relations or spatial scales, this will be at best piecemeal, and will not support any universal ordering or relation between scientific practices. Ditching a universal notion of levels correspondingly abandons traditional debates about reduction and emergence. Lastly, Potochnik argues that the independence of scientific fields and their lack of integration is not a negative. While we should not hope for integration between existing fields, Potochnik thinks that new fields can develop to fill gaps in our understanding. This patchwork of non-integrated but mutually supporting products points to a Neurathian conception of scientific unity.

Throughout, Potochnik develops her points with regards to an impressive array of examples – including sexual selection in flies, human aggression, population genetics, climate science, and obesity. The best aspect of the book is that it gives a global and generally appealing account of the combined metaphysics and epistemology of real science. Despite her focus on understanding, Potochnik is no subjectivist. Our practices are constrained and shaped by engagement with the world, and genuine understanding is based on actually uncovering causal patterns – although our representations of the phenomena they contribute to are bound to be piecemeal and idealized. The account, despite its psychologistic leanings, thus hopes to ground a limited scientific realism. “Aha” moments are not sufficient to account for genuine “grasp” of the phenomenon at issue. Nor, however, does the account posit unrealistic epistemic access to the world, which outstrips what we can reasonably say about our cognitive capacities. Attaining a balance between the extremes of subjectivism and unrealistic epistemic access is a central issue in the epistemology of science. Potochnik’s account is rich, original, and in general very convincing.

In such a rich account, as you would expect, there is much to quibble with, and plenty to disagree about. Some will take issue with Potochnik's decidedly liberal account of causal patterns. For Potochnik, what matters is patterns of dependence, not what *actual changes* contributed to the phenomenon. If an organism's foraging patterns *would* have been shaped differently by a different distribution of resources, for instance, that counts as a causal pattern even if resources were stable during the organism's actual history. Others will debate whether Potochnik's view is too psychologistic. I won't focus on these here, since I am in strong agreement with Potochnik on both.

Other parts of the book I found less satisfying. For instance, despite the focus on scientific products and reasoning, there is relatively little discussion of the nature of scientific representation. Potochnik is officially neutral on the representation relation, but often opts for a vague notion of similarity, one that could use considerable fleshing out (Weisberg, 2012). Second, Potochnik is too heavily focused, for my liking, on generality in pattern-explanations. Despite noting that generalizations are always limited, Potochnik argues that one of the main aims idealization serves is to increase generality in our represented patterns, and that "broad causal patterns are the path to human understanding" (144). I think that this is considerably overstated, since I believe that, at least in some contexts, the temporal, quantitative, and contrastive aspects of patterns can be explanatory independently of how widely they hold (Burnston, 2017), and in the importance of explanations even in singular cases (Bogen, 2005). Lastly, despite their centrality for the framework, there is no substantive discussion of how scientists individuate or characterize phenomena (Colaço, 2018).

In the space that remains, however, I will focus on two main issues. The first is the role that accuracy plays in the notion of "acceptability." The second is whether the division-of-labor model is the right kind of pluralism for describing scientific practice.

"Acceptability" is important for Potochnik – it is what helps the understanding-based view avoid collapsing into subjectivism. But Potochnik's phrasing of it depends on a notion of accuracy which is hard to cash out sufficiently given her other commitments. Accuracy is supposed to be weaker than truth, or even "approximate truth" (116), but we aren't given a positive characterization or any details about how scientists should measure accuracy. Instead, a representation is taken more accurate when it is "closer to the truth," and an idealization is acceptable when its "divergence from the truth is insignificant" (100). Moreover, individual research programs must assess "whether truth or accuracy in some regard is required" for their aims (192), and, in certain circumstances, idealizations "must accurately represent the causal factors contributing to [the focal] pattern" (157).

This is problematic, because on Potochnik's framework *all* representations idealize. They all falsely represent the phenomenon. If that is the case, then what independent rubric is there by which to assess *how* close to the truth the explanation is? Now, part of Potochnik's motivation here is to avoid saying that a representation must meet some universal standard of accuracy to count as explanatory. I agree as far as this goes, but one might want more guidance for practice from so central a part of Potochnik's view. I can think of two ways Potochnik might try to bulk up the notion of accuracy from within her account, but neither is particularly promising.

First, Potochnik says in many places that representations idealize away from potential causal factors *besides* the focal ones in the pattern. So, one might hope that the factors that *are* represented could be represented accurately, even if the system/phenomenon as a whole is represented in a falsifyingly simplistic way. However, this does not square with some of Potochnik's examples. For instance, she resists fictionalism by arguing that idealizations such as classical electron orbits can represent, if falsely, the key causal pattern of interest. Even her illustrative example of the San Francisco BART map falsely represents *the very relationships* it is meant to convey, by showing all stations as on straight-line routes and equidistant. On a view where idealization is so central and thorough-going, it is hard to pick out a place where an intuitive notion of accuracy fits.

Another avenue is to look at different ways that models are justified in practice, for instance by empirically corroborating their assumptions or checking their predictions against the data. However, these important aspects of practice don't seem to do much to solve the worry about accuracy's role. The main thrust of Potochnik's view is that representations can attain empirical success *despite* – indeed, partially *because of* – their falsifying tendencies. So relying on this success to establish their accuracy is not particularly forceful. Thus, more is needed in Potochnik's account of acceptability.

The second main point – and this is really the big one for me – is whether the division-of-labor model is either descriptively or normatively adequate as a view of scientific practice. Potochnik thinks that different representations will require different idealizations in focusing on their causal patterns. And she thinks scientists *should* be free to pursue these patterns, to whatever extent they can model them acceptably, without needing to give “full” explanations – i.e., without needing to account for other causal patterns in their explanations. This is not to say that integrations between models and patterns don't or should never occur. Sometimes, as in the case of environmental influences on gene expression in particular organisms, the causal thicket will simply be too intertwined to admit of separating patterns. In other cases, integrated fields may spring up to study focal patterns not represented well by extant subfields. This is Potochnik's account of how evolutionary developmental biology came about. Still, Potochnik thinks these kinds of examples are the exception not the rule. In most cases, scientific products and research programs are “steadfastly independent” (195), although they may sometimes generate evidence for each other.

In assessing the descriptive adequacy of the division-of-labor model, we can ask how well it extends to fields of major interest for philosophy of science. Consider systems biology and systems neuroscience. Ukai and Ueda (2010) argue that understanding a complex system like an intracellular circadian clock requires traditional bench science to identify the components of the mechanism, general design principles to conceptualize the organization of those components, and synthetic construction of the system to check the sufficiency of one's understanding. But these different approaches significantly overlap in what they explain – namely the patterns of interactions of molecular components that result in a sustained circadian oscillation. Similarly, in neuroscience, statistical and abstract computational analysis is often pursued in tandem with detailed mechanistic understanding of what particular brain areas are doing, and it is at least arguable that these both vitally contribute to explaining what the system does *in any given case*

(Burnston, 2016). I could multiply examples, but the point is that different empirical and modeling approaches in these contexts go far beyond just providing evidence for each other; they are explanatorily cooperative. This extends to the normative aspect as well – Bechtel and Abrahamsen (2010) have argued that our best way of understanding complex biological systems involves combining distinct approaches. This is not to say that the division-of-labor model is never right, but that its own scope may be limited.

Another potential problem with division-of-labor models is that, taken normatively, they risk letting particular fields off too easy in terms of drawing fences around what they care about. Take any major theoretical debate of your choosing – say, classical versus connectionist accounts of language learning, or adaptationist versus empiricist accounts of imitation (Heyes, 2018). Whenever there is a major theoretical conflict, is this simply due to frameworks having different focal patterns? For what it's worth, many evo-devo proponents are not so sanguine about its relationship to traditional genetic approaches. They think that traditional approaches get something wrong about the *very phenomenon* they purport to study by not taking developmental hypotheses into account, and thus giving “too much causal significance ... to genes and selection” (Laland et al., 2015, p. 6). In sum, any account of scientific pluralism must describe when and how scientific fields/products will productively cooperate, and when they will genuinely conflict. It is not obvious that the division-of-labor model, rather than more integrative forms of pluralism, accurately does so.

A book with which one had no disagreements would be an uninteresting read. Because I am so on board for Potochnik's general approach, and found so much of her picture incredibly appealing, the book provides an extremely useful touchstone for the points about which I do disagree. I expect that I will return to it frequently as I pursue my own projects for useful ideas, contrasting viewpoints, and helpful articulations of general principles. I can't think of a better endorsement to give to a philosophical text.

References

- Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3), 321-333.
- Bogen, J. (2005). Regularities and causality; generalizations and causal explanations. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 397-420.
- Burnston, D. C. (2016). Computational neuroscience and localized neural function. *Synthese*, 193(12), 3741-3762.
- Burnston, D. C. (2017). Real Patterns in Biological Explanation. *Philosophy of Science*, 84(5), 879-891. doi:10.1086/693964
- Colaço, D. (2018). Rip it up and start again: The rejection of a characterization of a phenomenon. *Studies in History and Philosophy of Science Part A*, 72, 32-40.
- Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*: Harvard University Press.

- Laland, K. N., Uller, T., Feldman, M. W., Sterelny, K., Muller, G. B., Moczek, A., . . . Odling-Smee, J. (2015). The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc Biol Sci*, 282(1813), 20151019. doi:10.1098/rspb.2015.1019
- Mitchell, S. D. (2002). Integrative pluralism. *Biology and Philosophy*, 17(1).
- Rice, C. (2015). Moving beyond causes: Optimality models and scientific explanation. *Noûs*, 49(3), 589-615.
- Ukai, H., & Ueda, H. R. (2010). Systems Biology of Mammalian Circadian Clocks. (November 2009), 1-25. doi:10.1146/annurev-physiol-073109-130051
- Weisberg, M. (2012). Getting Serious about Similarity. *Philosophy of Science*, 79(5), 785-794.