A Study of Explanatory Virtues

Daniel C. de Coimbra*, Marco A. C. Ruffino

Abstract

The inference to the best explanation (IBE) appeals to how hypotheses would explain certain facts. If a hypothesis explains well and better than any other (available) hypotheses, IBE concludes that it is true. Our work examines what makes it the case that some hypotheses would explain facts better than others. We also examine other non-explanatory features which are (thought to be) correlated with the truth of a proposed explanation. We name all these features *explanatory virtues*, and they are the measure employed by IBE of goodness of explanation. Our research has the purpose of investigating four chief (proposed) explanatory virtues: simplicity, unification, *non ad hocness*, and coherence. We intend to study the description and justification of these proposed virtues. Our final goal is to improve and appraise the employment of IBE, a valuable tool in philosophy

Key words:
Epistemology, explanation, simplicity.

Introduction

An explanation, we propose, is a *reason why* something is the case. For instance, causes are reasons why their effects obtain, while arranged parts are reasons why their wholes exist. Some proposed explanations fare better than others. Some would explain more, while others are more likely to be true due to other reasons. The features relevant for these distinctions we call explanatory virtues. Total explanatory virtue is a composite property, combining features such as relevance, informativity, unification, generality, coherence with our background knowledge, lack of *ad hoc* accommodations, and sometimes even simplicity, elegance, and, in the context of physics, naturalness and non-fine-tuning. Explanatory virtue is thought to be indicative that the theory provides a true and complete explanation of something. Our research seeks for the best description of four explanatory virtues (simplicity, unification, *non ad hocness*, and coherence) and asks whether or not they are truth-conducive. That is, are explanatory theories with these characteristics more often true?

Results and Discussion

This is a preliminary presentation. Our research is still ongoing. Here we report promising leads rather than conclusions. The methodology of philosophy is assessing descriptive theses by rational argument rather than by empirical testing. This does not mean that no use is made of antecedently-obtained empirical information. Here is the kind of information required to appraise each virtue. Simplicity requires the statistics of curve-fitting, complex systems theory, and evidence about simplicity in nature from empirical science. However, there is an *a priori* argument that has caught our attention: (1) If simplicity is ignored, evidence underdetermines theory. (2) We rely on simplicity and are often pragmatically successful. (3) Therefore, our choice of the correct theory is partially due to our attention to simplicity. Moving on, *non ad hocness* requires historical and psychological information about theory formation. There's reason to believe *ad hoc* theories often contain "fudging" factors not transparent to their creators, but which would be weeded out by tests of novel prediction (which remove *ad*

hocness). These factors make it likely that the theory is false. Yet, when the presence of these factors is likely to be transparent, *ad hocness* may be unproblematic. Finally, coherence and unification can be assessed with information about the coherence and unification of explanations (*reasons why*) in reality, a difficult task perhaps best pursued in the metaphysics of physics and in applied systems theory. We have yet to pursue this line of work. We are pessimistic about philosophy being up to the task of discovering the truth about these matters. We are hopeful that we will at least obtain clarity about the issues involved, its connections to other problems, and the limitations of our current knowledge (at least among non-scientifically-savvy philosophers).

Conclusions

The past successes of inferences to the best explanation creates hope, but perhaps the success was due to the domain-specific wisdom of its users. It is possible that no general defense of the above explanatory virtues is forthcoming. Simplicity and unification appeals may be effective only in certain ways and in certain domains, as certain aspects of reality may be complex and fragmented. Coherence and *non ad hocness* seem the most likely to gain uniform support, but these too suffer from certain difficulties. Overall, the inferential strength of IBE is not yet clear.

Acknowledgement

_____

[1] Lipton, P. *Inference to the Best Explanation*. 2. ed. London: Routledge, **2004**.
[2] Skow, B. *Reasons Why*. Oxford: Oxford University, Press, **2016**.
[3] Psillos, S. *Knowing the Structure of Nature:* Essays on Realism and Explanation. Hampshire, UK: Palgrave Macmillan, **2009**.
[4] Thagard, P. "The best explanation: Criteria for theory choice." *Journal of Philosophy*, **1978**, v. 75, n. 2, pp. 76-92.
[5] Cartwright, N. *How the Laws of Physics Lie*. Oxford: Clarendon Press, **1983**.
[6] Sober, E. *Ockham's Razor*. Cambridge: Cambridge University Press, **2015**.