

## NOTES PHILOSOPHIQUES

# EXPLANATION AND UNDERSTANDING REVISITED

ANDREI IONUȚ MĂRĂȘOIU

**Abstract.** Intuitively, it is often claimed, when we understand something, we can explain why it occurs the way it does, at least in outline. This piece of common sense is often encountered in contemporary epistemologies of understanding. Against this view, Peter Lipton distinguishes explanations from their cognitive benefits. He argues that those benefits, which lead to understanding the phenomena explained, can be obtained without recourse to explanations as well. It would follow that we can understand without being able to explain the phenomena understood. My appraisal of this debate is twofold. First, I reply, point by point, to Lipton's objection, defending (for the purposes of argument) the view that we can only understand what we can explain. Secondly, however, I believe that this does not fully support the view that the ability to explain what one understands is a prerequisite for understanding it. Rather, I argue that differentiating between explanations (as verbal end-products of explaining), the process of explaining itself, and the cognitive benefits one may reap from understanding and/or explanation amounts to a moot distinction, which can be recast in a pluralist vein. This, I argue, questions whether the debate about how understanding and explanation relate is ultimately substantive or, at least in large part, a semantic issue concerning how we choose to devise or use key epistemological terms. I illustrate my discussion of Lipton throughout with reference to a specific model, Volterra's equations governing the co-dependence of prey and predator species in a habitat.

**Keywords:** scientific explanation, understanding, the Lotka-Volterra model.

## 1. INTRODUCTION

The exploration to follow aims at questioning whether the links posited between understanding and explanation are metaphysically substantive or if they are, on the contrary, offshoots of semantic decisions concerning how to best use, in many but perhaps not all contexts, the *words* "explanation", "understanding", and their respective cognates.

Sweeping pronouncements to the effect that a debate is overtly metaphysical or covertly semantic make little headway. So, rather than making any such

Andrei Ionuț Mărășoiu ✉  
University of Bucharest

pronouncement, I illustrate the concern that animates this project with a traditional debate in the epistemology of understanding, namely, whether we could ever understand without having ready to hand, or having available (in some way), an explanation of the target phenomena or events thus understood.

Peter Lipton, in his 2009 chapter “Understanding without Explanation”, has argued that we can understand and enjoy the cognitive benefits of explanation without necessarily having such an explanation of what is understood immediately available. I reply to Lipton’s discussion point by point, illustrating both his remarks and my rejoinders with one important model in population ecology, the Lotka-Volterra model. My aim in doing so, however, is not to support a metaphysically inflationary view of explanation as a prerequisite of understanding (epistemically and cognitively), although much of my reply to Lipton (bracketing my own tentative conclusions and running commentary and focusing on the dialectic alone) could be read that way.

Rather, what I aim to suggest is that even though Lipton’s project is quite interesting metaepistemologically, replies in defense of the view he criticizes seem – and perhaps always so – to be readily available. This, I suggest, though I propose no watertight argument in favor of my suggestion, might tentatively make us wonder whether debates that aim to ultimately ascertain if explanation is a prerequisite for understanding are substantive debates or if, on the contrary, they are purely semantic or pragmatic, concerning how we choose to use such categories in typical cases. If we entertain the second theoretical possibility, then the issue no longer concerns our cognitive ability to explain phenomena we think of, or what it might cognitively mean to understand those very phenomena, but rather how we, as a community, choose to use words like “explanation”, “understanding” and their cognates on those occasions where such uses might be salient or illuminating.

## 2. PREREQUISITES FOR UNDERSTANDING

The literature on understanding in epistemology and general philosophy of science is rife with identifying putatively necessary conditions for understanding – or with challenging such identifications. In this section, I indicate a few of the more frequent candidates advanced as prerequisites for understanding.

One of the most robust intuitions concerning understanding is that it often, perhaps always, involves being able to explain what one understands. Several authors have picked up on the theme. Thus, Zagzebski (2001, p. 240) writes:

Understanding is a cognitive state that arises from *techne* ... The person who has mastered a *techne* understands the nature of the product of the *techne* and is able to explain it. She also knows the good in a sense that gives her a common understanding with the practitioners of other *technai*.

And Grimm (2006, p. 517): “it seems to be a hallmark of having understanding that we can typically articulate (or explain) what it is in virtue of which we take ourselves to understand”. Strevens (2013, p. 513) also distinguishes understanding-why from understanding-that; it would then be almost banally true that understanding-why corresponds with explanations, which answer why-questions.

To say that understanding is based on *skills and know-how*, or that it requires being able to *explain* what one understands, or that it is *factive*, or that it is partly constituted by *knowledge*, is only a small list of epistemic values the realization of which has been advanced as required for understanding. Realizing these values is not, at least *a priori*, mutually exclusive. They can be realized jointly. Indeed, they might, in some cases, all be realized, e.g., if understanding partly requires having explanatory abilities that produce knowledge that answers why-questions.

However, the bare mention of such putative prerequisites invites the question of whether counterexamples exist: whether we might sometimes understand without realizing any (or, perhaps, all) of these epistemic values. Lipton’s arguments in favor of the existence of “understanding without explanation” belongs to this category. I now turn to his criticism.

### 3. EXPLAINING AND ITS BENEFITS

Lipton’s (2009) “Understanding without explanation” produces several classes of examples where he argues understanding can be had in the absence of explanations. Lipton’s general strategy is to identify the kinds of cognitive benefits explanations usually bring, to argue that those benefits are what constitutes understanding, and then to support the claim that the benefits might perhaps be gotten in ways which don’t involve explanation at all. It would follow that understanding can be had in the absence of explanation. As he puts it, “by distinguishing explanations from the understanding they provide, we make room for the possibility that understanding may also arise in other ways” (Lipton, 2009, p. 44). Lipton discusses scientific understanding alone, and prudently stays neutral about whether his remarks extend to other kinds of understanding, and his restriction will be presupposed throughout in what follows. I only tentatively submit that Lipton’s initial view might perhaps be generalized.

While I myself don’t think this amounts to positive support for the view that being able to explain what one understand is a prerequisite for understanding it, in what follows I will dispute Lipton’s strategy point by point. I agree with him that identifying the cognitive markers of understanding is central to the project of clarifying what understanding is. However, unlike Lipton, I think we may approach this debate from a perspective in which understanding and explaining go hand in hand, even though we may also approach it from Lipton’s alternative perspective as well. I will review the cognitive benefits Lipton invokes, and argue they don’t present counterexamples to the claim that understanding is a kind of explaining.

One item on Lipton's list of cognitive benefits of explanations I leave aside. I won't discuss feelings of understanding. As Lipton remarks, our understanding something may be manifested in feelings of surprise, expectation, or familiarity. He rightly points out that these feelings can be had in the absence of being able to explain phenomena one approaches. However, Lipton doesn't argue that these feelings are had *rationally* – thereby manifesting understanding – when explaining is indeed absent. Lacking the central epistemic dimension of rationality in their description, it is difficult to see how feelings of familiarity (and other such metacognitive feelings) could sever ties between understanding and explanation, two centrally epistemic types of higher cognitive states<sup>1</sup>.

For concreteness, I will focus discussion on a specific model from ecology, the Volterra equations for a habitat with a prey and a predator population. This should help sharpen the issues, and show how general points about understanding can be applied to concrete cases in the history of science.

But the main point I wish to make is fully general, and can be abstracted away from my discussion of Volterra's model. It is this: Whether understanding is a kind of explaining or not depends on how we conceive of both. Lipton (2009, p. 43) writes:

Explaining why and understanding why are closely connected. Indeed, it is tempting to identify understanding with having an explanation. Explanations are answers to why questions, and understanding, it seems, is simply having those answers. Equating understanding with explanation is also attractive from an analytic point of view, since an explanation is understanding incarnate. The explanation is propositional and explicit. It is also conveniently argument shaped, if we take the premise to be the explanation proper and the conclusion a description of the phenomenon that is being explained. So we are on the way to specifying the logic of understanding. The explanation is propositional and explicit. It is also conveniently argument shaped, if we take the premise to be the explanation proper and the conclusion a description of the phenomenon that is being explained.

This is how Lipton starts his text, and we already see a sliding from “Explaining why and understanding why are closely connected” to “Equating understanding with explanation”. As Lipton puts it, “The explanation is propositional and explicit.” There is something amiss with Lipton's picture of explanation. Here is Strevens (2013, p. 512) commenting:

Against Lipton, I suggest that understanding that, or grasping, a proposition does not imply an ability to make the proposition explicit. I can grasp a correct explanation, then, that I am unable articulate... I might add that a particular individual's inability to communicate the explanation they grasp does not imply

<sup>1</sup> For more on the role of conscious experience in understanding, the differing viewpoints of Trout (2002) and de Regt (2004) are instructive.

that the object of their grasping is inherently incommunicable or ungiveable, but merely that they are presently the wrong person for the job.

Strevens seems largely right, though he leaves it mysterious what it might be for an individual to grasp a propositional explanation without being able to verbalize it. The solution to this particular aspect of the debate at hand, I suggest, is to distinguish being able to explain something from its propositionally articulated results. We are able to answer why-questions often enough *ex tempore*, even *without* invoking some already conceptually articulated belief we will have stored in memory. Often, answering why-questions is a process of *articulating* such answers, not of retrieving them already packaged from memory. Explaining differs from being in possession of an explanation.

One of Lipton's reasons why understanding differs from explaining is that, if we suppose Kuhnian historiography of science is on the right track, it seems that case-based reasoning (or paradigmatic problem-solving in normal science) gives us understanding while not delivering any propositionally articulate explanations. This justification breaks down once we make the distinction between being able to explain something and being in possession of a ready-made explanation. Either normal scientists who don't yet have verbal explanations are *able* to explain the phenomena in question, so we can recognize their understanding of those phenomena. Or they are unable to produce such explanations. If the latter, then it is not the mere absence of explanations, but their inability to produce any should undermine the claim that they genuinely understand the phenomena they approach – as opposed to, say, swiftly but irreflectively solving problems which involve them. Either way, understanding and explaining have not yet, or not thereby, been shown to part ways.

Granted, understanding – especially when tacit and model-based – may not always be accompanied by ready-made explanations. However, Lipton does not seem to have shown, as I will argue, that understanding cannot always partly consist in our ability to produce explanations. I'll now illustrate how this discussion applies to Volterra's model.

#### 4. VOLTERRA'S EQUATIONS

In my discussion of how understanding relates to explaining, I will follow closely Volterra's (1927/1978, pp. 80–89) presentation<sup>2</sup>. The equations I am interested in characterize the evolution of a predatory population ( $N_2$ ) and a population it preys upon ( $N_1$ ). Other equations investigated by both Lotka and Volterra concern the evolution in tandem of populations sharing a food supply,

<sup>2</sup> Andersen (2017) gives a presentation in many ways clearer. I choose this particular presentation by Volterra because its mathematics is elementary and intuitive: any student of freshman calculus can follow it. Andersen's presentation is more traditional, and more in the spirit of Volterra's work, but it loses a bit in simplicity.

co-prey populations, prey-predatory proportionalities generalized for the case of  $n$  populations. I stay with the two-species case because it seems to me that the conceptual points that can be made with those refinements of the model can also be made by using this simpler but more abstract version.

In the two-species case, the equations codify the intuition that population sizes of prey and predatory species depend on one another. The rate at which the preyed-upon population grows ( $\frac{dN_1}{dt}$ ) depends on how large the predatory population ( $N_2$ ) is; and the rate at which the predatory population grows ( $\frac{dN_2}{dt}$ ) depends on how large the prey population ( $N_1$ ) is.<sup>3</sup>

The equations are:  $\frac{dN_1}{dt} = (\varepsilon_1 - \gamma_1 N_2)N_1$  and  $\frac{dN_2}{dt} = (-\varepsilon_2 + \gamma_2 N_1)N_2$ . If no predators existed, the preyed-upon population  $N_1$  would grow at a rate of  $\varepsilon_1$ . If no preys existed, the predator population  $N_2$  would decrease at a rate of  $\varepsilon_2$ . With predators in its habitat, the preyed-upon population will increase at a slower pace, and how slowly is a rate  $\gamma_1$  of how many predators there are ( $N_2$ ). With preys in its habitat, the predatory population will decrease at a slower pace, and how slowly is a rate  $\gamma_2$  of how many preys there are ( $N_1$ ). Volterra's (1927/1978, p. 80) metaphor is apt: " $\gamma_1$  measures the susceptibility of species 1 to predation and  $\gamma_2$  measures the predatory ability of species 2."

The Greek might be easier if illustrated with the tale of foxes and hares. If there are plenty of hares in a habitat, foxes will have a richer food supply and so their numbers will grow. With more foxes around to feed, hares will nearly run out, which will in turn produce a drop in the numbers of foxes, resulting in a spike in the number of hares. Barring external interventions (other predators, other preys, human intervention, a change in climate, etc. – notice this is quite a sizeable idealization), the cycle can perpetuate itself.

With the bare bones of Volterra's model laid out, I'll now illustrate Lipton's discussion with it.

## 5. TACIT UNDERSTANDING

One of the most important red threads running through Lipton's text is the idea that understanding may often be tacit – unexpressed, and perhaps inexpressibly fully, in propositional explanations. Such tacit understanding may be revealed by scientists' skills, by their use of models that cannot be fully captured in words, and by the many fictional or instrumental uses of models, hypotheses and background theories, which have no explanation in view or in the offing. He writes:

<sup>3</sup> Changes in a magnitude  $f$  are measured by the first derivative of  $f$  with respect to time  $t$  ( $\frac{df}{dt}$ ) if one exists, because change occurs in a unit of time. The parenthetical "if one exists" is important. Only continuous functions have derivatives, and a population size is a discrete variable. Volterra's equations presuppose that we treat discrete variables as continuous. This is an idealization met often: there are no half-births or deaths.

As soon as we consider the possibility that causal information may be the content of tacit knowledge, the prospect for routes to understanding that do not pass through explanations becomes bright, since causal explanation, I suggest, requires that the information be given an explicit representation. In short, there is such a thing as tacit understanding, but not tacit explanation (Lipton, 2009, p. 45)

It is worth exploring Lipton's view in the context of a real-world scientific product, Volterra's model just presented. Someone wishing to challenge Lipton has two routes to go.

Here is the first one. It is always possible to think that, for example, Volterra's model lacks explanatory power, that it merely models changes in prey-predator ratios which it fails to illuminate. Correspondingly, it is always possible to think that Volterra's model provides us with no understanding of how preys and predators interact. If none obtain, then understanding and explaining relate in a banal way. This possibility is always in the background, and should be considered. On this reading, it is still always possible to insist that explaining and understanding go hand in hand. The possibility that neither explaining nor understanding based on this model are possible will always be in the background. However, in what follows, I'll primarily consider a second route, namely, that Volterra's model is *both* explanatory and conducive to understanding – each, to some degree or other.

The challenge Lipton raises is that we sometimes lack explanations yet possess understanding. To illustrate with Volterra's model: What do a pair of equations have to *say* about the natural world? Nothing at all, at first blush; they *seem* explanatorily blind. The equations only model the numerical dependencies between the population sizes of predators and preys, e.g., moose and wolves (Andersen, 2017). However, *contra* Lipton, from the fact that Volterra's equations *alone* don't explain anything – which I grant – it doesn't follow that they aren't part of good (potential) explanations. Lipton anticipates this reply, writing:

One might say that, even if the causal knowledge gained in these cases is partially tacit, the knower will nevertheless be able to say something, however oblique, about the causes. This may well be so, but what the knower can say in these cases does not exhaust what the knower understands. The cases could also be resisted by either expanding our notion of explanation or contracting our notion of understanding. But neither maneuver is plausible here. A tacit explanation would be an explanation one could not give, even to oneself, and this does violence to the notion of explanation. And to deny that this tacit knowledge constitutes understanding would be Procrustean. We want a way of marking the difference between someone who knows that the phenomenon occurs but has no inkling why, and someone who has a deep and subtle appreciation of its causes, though he or she is not in a position to articulate that knowledge. "Understanding" seems to be the right word for the difference. It is not *ad hoc* to insist that explanations must be verbal but that understanding need not be. (Lipton, 2009, p. 46)

But it is *ad hoc*, I should say. The practice of explaining something is a discursive practice with cognitive roots. Few practices are ever *identified* with their end-products: explanations, or answers to why-questions. In answering a why-question, one doesn't retrieve and recite an answer learned by rote. Rather, one *articulates* an answer by typically thinking it *through* as one says it. Thoughtful answers aren't ready-made.

To illustrate: We expect the following. Predators will grow populous until prey starts running out, then predators will start running out, followed by a comeback of prey, followed by a comeback of predators. (Recall the foxes and hares cartoonish example above.) That is, we expect a cycle governing how the sizes of prey and predatory populations depend on each other. With a bit of elementary mathematics, Volterra (1927/1978, p. 89) vindicates that commonsensical idea, and concludes with a beautiful result, namely, that the period of the cycle approximates  $T = \frac{(2\pi)}{\sqrt{\varepsilon_1\varepsilon_2}}$ .

Andersen (2017) argues, I believe persuasively, that there is not one, but two ways in which the result just mentioned can be explanatory. One route is to consider Volterra as giving the bare bones of an empirical model: supply parameters  $\varepsilon_1$ ,  $\varepsilon_2$ ,  $\gamma_1$  and  $\gamma_2$  for particular pairs of predator-prey populations, and an empirical model results. If so, the fact that no two populations are ever in equilibrium – when considered in isolation from the rest of their ecosystem – is straightforwardly explained by Volterra's equations. It is explained *by* the model. That is, having the model *enables* any competent user to *produce* the requisite propositional explanation.

A second route Andersen (2017) points to is mathematical. Here, one doesn't consider pairs of populations and ask whether the equations fit them. Rather, one shops from the get-go for pairs of populations that fit these equations. For some pairs of populations, the lack of a steady state (or static equilibrium) will be no empirical chance. It will be a mathematical result, flowing directly from Volterra's model. This is because those pairs were *so chosen* so that the equations do or do not fit them. Again, what this means is that competent users of the model are enabled – are in a position, other things being equal – to explain why a steady state cannot be reached, and why the co-development of a prey and predatory population considered in isolation is cyclical.

On both construals, models *themselves* are explanatory – be it an empirical or a mathematical model – because competent users may engage in explaining, even if they don't have ready-made explanatory propositions ready to hand. Cyclicity of predator-prey ratios is nowhere *verbally* explicit in Volterra's equations or in resulting empirical models. This is *tacit* knowledge, our know-how of using a mathematical or an applied model to answer why-questions about preys and predators.

Before leaving this section, I would like to indicate what I think Lipton's contributions makes crystal clear (perhaps despite his own leanings); I will return to the issue in concluding the text. Recall his words: “The cases could also be resisted by either expanding our notion of explanation or contracting our notion of

understanding. But neither maneuver is plausible here” (quoted above). This, I believe, is exactly right. A metaepistemology that is flexible enough as to afford our varied and context-sensitive knowledge practices needs to consider these alternative construals, stretching the notions of explanation and understanding respectively so that they may, or may not, fit, according to how we then and there choose to substantiate them. If these remarks by Lipton are on the right track, then we might make it seem as though explanation and understanding match; and we might also make semblance of the opposite. The deeper point being that such appearances might deceive in that notions of explanation and understanding, respectively, are flexible enough as to afford both kinds of approaches.

## 6. VARIETIES OF NECESSITY

Lipton (2009, pp. 46–47) writes:

there are other forms of understanding that suggest routes that are propositional and explicit but still do not involve explanations. Consider the necessity conception of understanding (Glymour, 1980). On this view, to understand why something occurs is to see that, in some sense, it had to occur.

What are some cases illustrating it?

The cases I seek are arguments that are not explanations but do generate understanding by showing necessity. There are a number of plausible candidates, including various symmetry and optimization arguments in physics. Thought experiments are another promising source.

Here, the thought seems to be that ascertaining necessity makes other explanations redundant or useless – since any other explanation could only explain why something happened, not why it had to happen. Whereas, if it had to happen, nothing more needs to be said about why it did indeed occur. Are there such special cases of understanding? Lipton goes on to illustrate:

For example, Galileo’s thought experiment that demonstrates that gravitational acceleration is independent of mass gives one an understanding of why masses must fall with the same acceleration even though his experiment may not offer an explanation for that phenomenon. Suppose that heavier things accelerate faster than light things, and consider a heavy and a light mass connected by a rope. Considered as two masses, the lighter one should slow down the heavier one, so the system should accelerate slower than the heavier mass alone. But considered as one mass, the system is of course heavier than the heavy mass alone and so should accelerate faster than the heavier mass alone. But the system cannot accelerate both slower and faster, so acceleration must be independent of mass (Lipton, 2009, p. 47)

One way to react to this in keeping with the view that understanding is a kind of explaining is to construe logical reasoning as explanatory because it proves the independence of two worldly properties, mass and gravitational acceleration. Another way would be to question that this argument provides genuine understanding. Yes, the argument demonstratively settles the issue of the independence of two properties. But we thus come to know *that* they are independent, with no inkling of *why* that's the case. This, in a sense, is precisely Lipton's point; only he thinks we do gain some understanding. I doubt that, because logical necessity holds *however* the world is arranged. Here is Strevens' diagnosis:

The Galilean argument indeed provides, I think, a sense of illumination. It does so by harnessing a certain physical intuition, that the presence of the rope is not the sort of thing that influences the dynamics of falling. The illumination depends, in other words, on the listener's pretheoretical grasp of physics. But it amounts to genuine understanding *why*, I suggest, only insofar as the psychologically operative pretheoretical physical principles constitute a part of the correct physical explanation of the independence of acceleration and mass; otherwise, the listener is like the young earth creationists..., apprehending an explanation, perhaps, but not the right one. My contention, then, is as follows: either the physical intuitions invoked by the Galilean argument are genuinely explanatory or they are not. If they are, then there is understanding *why* in virtue of a grasping, in this case implicit, of a correct explanation (or rather, a part of such an explanation); if they are not then there is no understanding *why* but only the impression thereof. (Strevens, 2013, p. 514)

Again, notice that Strevens' reply amounts to asserting what Lipton had formulated but denied: "The cases could also be resisted by either expanding our notion of explanation or contracting our notion of understanding." (quoted above).

To circle back to the same point using our target model rather than fix the Galilean example, let's see how talk of necessity might apply to Volterra's equations. As Andersen (2017) argues, Volterra's equations have a double aspect: conceptual and empirical. They are conceptual because they characterize the interdependence of two quantities in ways that are fully general. They are empirical to the extent that they afford empirical models, and they do: any pair of predator-prey populations which can be studied in isolation from their ecosystem.

This contrast affords a more nuanced approach to Lipton's suggestion that understanding, when it relies on necessities, doesn't always come in the wake of explanations. We can improve on that suggestion by noting that we should distinguish *three kinds of necessities*. The first is purely logical or mathematical: this, Galileo's argument and Volterra's equations share. The second belongs to applied mathematics: Galileo applies a *reductio* form of reasoning to some Aristotelian assumptions; Volterra applies his equations to predator and prey populations.

The third is nomological: Volterra's equations represent regularities *in nature*<sup>4</sup>. (Let me stress this is by no means intended as a contribution in the philosophy of logic. The classification based on modal strength suggested here is *pro tempore*, pragmatic, and for the purposes of this particular discussion. Many other classifications of modalities and their relative strength are possible.)

Once we do distinguish between kinds of necessities, we can make Lipton's suggestions more precise. It is quite possible, the thought goes, to have regularities – even ones holding with some form of necessity, be it conceptual or nomological – that don't come packaged with an explanation of why they hold. They are, as it were, *brute* regularities. Lipton may be suggesting they nonetheless afford some understanding. This way of putting things, however, raises questions. If regularities are brute, why think they give rise to understanding when they don't to explanations? And, conversely, what's the matter with explaining specific variations in, say, how moose populations varies with wolf populations, in terms of a regularity we otherwise might be inclined to think is brute? Explanations may sometimes work well enough even when they appeal to unexplained premises. Otherwise we would have to source everything down to first principles, which is impossible to achieve.

I conclude that Lipton hasn't successfully made the case that we may understand when we come to know of a necessity. The contrast between the Volterra equations and Galileo's *reductio* supports the following idea: not always, when we can make an assertion with a stronger modal force (nomological, mathematical, logical), does that imply there is more explanatory power. So far, Lipton would agree. But he would suggest understanding *does* co-vary with how strongly we can assert something, with its modal robustness. We haven't yet been given a reason for that. And the mere fact that we can use the phrase "I understand that" for Galileo's proof about mass and velocity doesn't imply that genuine understanding-*why* is procured by it, as opposed to mere knowledge of (modally strong) facts.

## 6. SEEING POSSIBILITIES

Lipton's appeal to modal understanding has another side to it. He also considers cases of models which provide understanding but no explanation because they explain how something *might* occur, even when it doesn't. As he puts it, "we can gain actual understanding from merely potential explanation" (Lipton, 2009, p. 47). The relevant explained possibility "may provide improved understanding of actual capacities, powers, and causal relations" (*idem*, p. 51). How does this threaten the idea that understanding is a species of explaining?

Since now explanation of one thing provides understanding of another, a counterfactual explanation of how the phenomenon could have come about (but did not) provides understanding of how it actually happened. (*idem*, p. 52)

<sup>4</sup> I leave the question of whether those regularities are laws aside here.

This is unpersuasive. After all, explaining a possibility is conducive to understanding because it shows that a law-like regularity fails to obtain, and that is useful information in articulating how we understand the phenomenon in question. The absence of a regularity is an actual state of affairs, not a merely possible one. And explaining why the contrary possibility obtained has immediate consequences about which regularities apply and which don't. Lipton hasn't offered a case where explaining and understanding come apart.

One might reply that all this is too quick, and that Lipton is after something: there is conceptual room between saying one explains a *possibility*, and saying one gives a possible explanation of what really *is* happening. So, Lipton thinks, partisans of the view that understanding is a kind of explaining have to misplace understanding – as the understanding of possibilities rather than of worldly phenomena.

Again, Volterra's model is a case in point to illustrate this how-possibly challenge. Given its highly idealized conditions (only two species; unless we wish to pack other influences into  $\varepsilon_1$  and  $\varepsilon_2$ ), the equations seem to characterize how prey and predatory populations *would* depend on each other if the *idealized* conditions obtained. For populations out in the wild, the challenge might go, equations explain nothing at all.

But one has to start somewhere. Whenever a problem is not tractable, one way to go is to find a simpler problem in the vicinity, making as few idealizations as possible – compatible with a swift solution. To illustrate this, let's revisit a concrete question raised above. Volterra raised the question of when an equilibrium state might be reached in a habitat with one population preying on another. We have seen, in the population cycle, that a large prey population invites a large population of predators, which then shrink the size of the preyed-upon population, which then shrinks the size of the predatory population. So prey and predatory populations can both co-vary and vary inversely, according to where we are in the cycle. If a "steady state" (Volterra, p. 81) is reached, it has to be reached when both populations stay constant, i.e.,  $\frac{dN_1}{dt} = \frac{dN_2}{dt} = 0$ . By plugging this condition into the initial equations, we get the steady-state population sizes of  $K_1 = \frac{\varepsilon_2}{\gamma_2}$  for the preyed-upon species and  $K_2 = \frac{\varepsilon_1}{\gamma_1}$  for the predatory species. Now *this* equilibrium state is *never* reached in the population cycle (Volterra, fig. 3, p. 87). It would be a miracle if, for every eaten prey, another one would be born.

However, in asking whether equilibrium is possible, we have to consider the fact that model-based inference supports – to an equal extent, if at all – conclusions *both* about possibilities and about the world. When it comes to equilibrium models and their explanatory power, we don't seem to have a discrepancy between understanding and explaining. According to Volterra's model, a steady state (where exactly as many preys are born as the predators eat) cannot be reached, provided its idealizations are assumed; that follows from a general account of a 2-species interaction. Volterra might be wrong, either in the equations, or in the idealizing

assumptions. We would then lack both genuine explanation and genuine understanding – since the model is wrong, in ways that fail to capture both the actual world and a large class of possible worlds where predator-prey equilibria obtain. Either way, however, no asymmetry obtains between understanding and explaining.

At the risk of overemphasis, let's return to what I will call, despite his own disavowal, Lipton's *dictum*: "The cases could also be resisted by either expanding our notion of explanation or contracting our notion of understanding." (quoted above). What could possibly stop us from doing so? Why fix on rigid notions of either explanation or understanding or both so as to prevent that? It would seem as though the debate boils down less to metaphysics and more to the semantics of how we use words like "explanation", "understanding", and their cognates.

## 7. MODEL-BASED UNDERSTANDING

Section 5 pitted Lipton's criticisms of the idea that understanding is always explanation-based against Strevens' defense of that idea. It is then incumbent upon me to show how the view I advance differs from that proposed by Strevens. The starting point of my comments are his remarks (2013, p. 512):

many explanatory models in science contain idealizations; interpreted literally, these models are false. Scientists gain understanding, nevertheless, by grasping idealized models. A reason to abandon the simple view's requirement of explanatory correctness...? No; at most a reason to use the term *correct* rather than, say, *true* – as the simple view already does. Why? The apparent falsehood of some models does not stand in the way of explanation, but that of others – as the young earth creationist case shows – most certainly does. The former class of models are "correct" in a sense that the latter class are not. Their correctness cannot, of course, consist in their literal truth. But if idealizing explanation is governed by any standard at all, there is a translation manual that, for any idealized model and context of (re)production, picks out a set of propositions that state the facts about the world that must obtain if the model is to be explanatory, and in virtue of which the model is explanatory, if they do obtain. Call these putative facts the explanatory content of the model...I hold that idealizations that appear to make some false assumption about the world... in fact make true claims about difference-making.

Why does this matter? For several reasons. I start with my adhesion and follow suit with my balks. First, I *do* believe one cannot genuinely understand something in the absence of a model of what is being understood. Strevens' line is well taken because it banks on this fact; I have just illustrated varieties of understanding based on Volterra's equations. However, as I see it, dependence on models is a *pragmatic prerequisite* rather than an epistemic necessity condition for understanding. Consider a different case. As I write this and you read this, we are both breathing.

Yet breathing, even if it is presupposed by our scholarly activity, has nothing at all to do with it. It is a prerequisite: we couldn't engage in academia if we didn't breathe. But it is not *epistemically* called for: breathing doesn't aid, in a *sine qua non* fashion, in procuring knowledge, explanations or understanding. Model-dependence, while I agree is pragmatically needed in order to understand anything at all, is, to my mind but not Strevens', much like breathing: it is ubiquitous but harmless in understanding (or in failing to understand, for that matter).

Secondly, to the extent that the explanations are accurate, so, *prima facie* at least, ought the models they are based on be. Alternatively, inaccurate models provide potential explanations which *would* lead to genuine understanding *were* they to be true, other things equal. The merely possible truth of such models leads, other things equal, to merely possible understanding reliant on them (with the caveats and discussion in Section 6).

Strevens seems to be assuming (for reasons related to the factivity of understanding) that all idealizations underwriting understanding have to be Galilean, i.e., that all idealized models M can be turned (transformed) into models M' free of idealization, only the transformation banks on a translation manual identifying the explanatory content of M, content which is preserved by M'. While this is a large topic to delve into, let me briefly address the issue here with the following remarks. I depart from Strevens at least on two counts. First, I doubt we can explain away all idealizations relevant to understanding (as Galilean). No knock-down argument has been provided for this, to my knowledge. It would be an admirable article of faith, but I see little reason to endorse it.

Secondly, I doubt – and this may be the crucial disagreement – that what we understand corresponds to the explanatory as opposed to the literal content of the models used in understanding. The issue concerns the putative requirement that explanations should – to the extent that they are genuine – be true to the facts, or correct. This requirement forces Strevens to distinguish the explanatory and literal contents of models, to the extent that understanding afforded by those models is to count as “correct” or approximately accurate *modulo* the idealizations that can be paraphrased away.

The point I make is, I think, straightforward (albeit admittedly telegraphic). From (doubtful) claims about the models that underwrite how we understand something, such as predatory-prey relationships, nothing follows about how those models are *cognitively* represented. It may be – by sheer accident – that precisely those models which preserve explanatory content are the ones that underwrite how we understand the phenomenon modeled. But this is a purely speculative possibility. More plausibly, how we cognitively internalize, the models in question has nothing to do with the *in-principle* translatability of idealizing assumptions those models are based on. And understanding is, first and foremost, *cognitive*.

I have, so far, argued against particular ways of challenging substantial relationships between understanding and explanation. But one may wonder exactly

what those relationships are. After all, John Turri (2015) thinks understanding is the goal of explanation, whereas Stephen Grimm (2010) thinks explanation is the goal of understanding. Is there a way to account for these views? I believe so. On the view Strevens and I jointly offer (differences aside), there is no talk of aims. Rather, *constitutively*, we haven't been given a compelling reason to think that one could understand something without being able to explain it. Correlatively, we haven't been given a compelling reason to think that one could come to understand something without becoming able to explain it. The needed explanations are extracted from the models used. (But, again, on my view, unlike Strevens', the models in question are not cooked-up philosophical counterparts of genuine scientific models, seeking to paraphrase away idealizations and other simplifying assumptions. My view leaves open the extent to which understanding corresponds to the facts.)

Let me once more stress the prefix: "we haven't been given a compelling reason to think that". For what I advocate, unlike what metaphysicians of understanding like Strevens and Lipton alike hold, is that, semantically speaking, we can stretch such notions as explanation and understanding variably, suited to context, so that oftentimes they may (perspective-relative) be seen to match, and at other times (perspective-relative again, and articulating such perspectives bears the onus of proof that Lipton has not substantiated in this text), could perhaps be construed to differ (I ultimately remain agnostic about this, I only discuss a set of reasons advanced in favor of this and find it lacking). I believe there is an interesting exploratory project in the offing here, about how the semantics of "explanation", "understanding", and related words connects with the epistemologies of explanations, understanding, and so on. That project is left for another day, but the dialectical discussion for and against Lipton carried out in these pages does, I think, suggest the vein in which a philosophy of scientific practice and deference to (or at least sympathies towards) epistemic contextualism might jointly lean us to appraise specific episodes in the history of scientific understanding.

## 8. CONCLUSION

Peter Lipton has argued that understanding can be had without explanation. To the extent that he conceives explanations as ready-made verbalized propositions, he might be right. However, I have argued that the interesting relation is not that between model-based understanding and verbalized explanations (in English, German, or other natural languages), but rather that between understanding and the *ability* to produce such explanations – between understanding and *explaining*. I have illustrated how understanding accompanies explanatory power with Volterra's model. The verdict can always be that, in given circumstances, we lack both understanding and the ability to explain what is happening. Or it can be that we enjoy both. Either way, we haven't been given any good reason to think they break apart, and that understanding can't be had when one is unable to explain what one understands.

Where next? The relation between understanding and explaining needs clarification. Khalifa (2012) has argued that more recent debates concerning understanding merely recapitulate an earlier literature about scientific explanation. But, as he points out, the differences Lipton theorizes, if they obtained, would undermine that idea. What I've shown here is compatible with Khalifa's deflationary view, on which talk of understanding can be fruitfully replaced with talk of explanations. Yet I believe we can make twofold progress in approaching the understanding-explanation connection.

I myself am tempted by Lipton's *dictum* (as phrased earlier): "The cases could also be resisted by either expanding our notion of explanation or contracting our notion of understanding." (quoted above). If this is right even roughly, then it might be that the explanation-understanding debate is less metaphysically substantive as it is semantically intricate. No one stops us, the thought goes, from using these words in the same contexts. And no one stops us (barring whatever epistemologists say) from using them differently. It then becomes an interesting sociolinguistic fact that we do tend to use words like "explanation", "understanding" and their cognates in more or less the same contexts. This pragmatic fact calls for semantic explanation. And here I can only conjecture, tentatively, that this tells tales of a connection presupposed between one's ability to explain phenomena and one's understanding of those same phenomena. True, our everyday habits of thought may well accompany our linguistic behavior, and they seem to more or less match the philosophy of different scientific practices and traditions. (Scientists of various stripes are, after all, *bona fide* members of their linguistic communities.) Still, nothing follows from this about the *metaphysics* of explanation, *or* of understanding, *or* of how they connect.

You might well dispute this. Perhaps there is a viable, substantive and important project about how understanding and explanation relate out there. If so, and if this is how you read Khalifa's, Strevens' and Lipton's remarks, then there will be, for you, something genuinely new and important which the contemporary literature about explanation-understanding connections tries to get at. The royal way to show that would be to identify what makes understanding *special*: why some explanations *fail* to deliver understanding, and which *extra* conditions might need to be met by one's ability to explain a phenomenon if one is to count as understanding it. Perhaps *problem-solving* usefulness might be a needed extra condition; or perhaps a particular set of skills, and the *expertise* one acquires throughout a career in research, might be importantly irreducible to properties of the models and theories one produces in research. I remain neutral among such options. But here lies a field of epistemic values that needs exploring. I would only invite such projects to anchor into specific episodes in the history of ideas (scientific, technological, or culture more broadly), and to pay heed to the bewildering variety of practices that underwrite ascriptions of both understanding and explanatory power based on similarly richly varied models built and used by those who thus come to enjoy both understanding and the ability to explain what they understand.

**Acknowledgments:** Writing the final version of this text was supported by “The effects of LLM interaction in digital and virtual environments on TOM” research project, funded by the University of Bucharest as a grant for enhancing institutional performance, no. 10504/25.10.2024. Writing the initial, quite different, version of the text, around 2016, was supported by the Jefferson Scholars Foundation through a John S. Lillard fellowship, and was one of the texts that surrounded my doctoral dissertation. I gratefully acknowledge the many hours of feedback received from my committee at the University of Virginia: Brie Gertler, James Cargile and the late Paul Humphreys. Many thanks also to audience members at the “Romanian-Polish Workshop in Causation, Understanding, and Knowledge: New Perspectives and Challenges”, jointly organized by the Romanian Academy, the Polish Academy of Sciences and the University of Bucharest, held on November 20<sup>th</sup>, 2023, where I delivered this talk; impetus from Mircea Dumitru, Sandra Brânzaru, Richard David-Rus and Paula Tomi was particularly helpful.

## REFERENCES

- H. Andersen, “Complements, Not Competitors: Causal and Mathematical Explanations”, in *British Journal for the Philosophy of Science* 69 (2), 2017, pp. 485–508.
- H.W. de Regt, “Discussion note: Making sense of understanding”, in *Philosophy of Science* 71 (1), 2004, pp. 98–109.
- C. Glymour, “Explanations, tests, unity, and necessity”, in *Nôus* 14, 1980, pp. 31–50.
- St. R. Grimm, “Is Understanding a Species of Knowledge?”, in *British Journal for the Philosophy of Science* 57, 2006, pp. 515–535.
- St. R. Grimm, “The Goal of Explanation”, in *Studies in the History and Philosophy of Science* 41, 2010, pp. 337–344.
- K. Khalifa, “Inaugurating Understanding or Repackaging Explanation?”, in *Philosophy of Science* 79, 2012, pp. 15–37.
- Th. S. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, 1962/2012.
- P. Lipton, “Understanding without Explanation”, in H.W. de Regt, S. Leonelli, K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*, University of Pittsburgh Press, 2009, pp. 43–63.
- J.D. Trout, “Scientific explanation and the sense of understanding”, in *Philosophy of Science* 69 (2), 2002, pp. 212–233.
- J. Turri, “Understanding and the Norm of Explanation”, in *Philosophia* 43, 2015, pp. 1171–1175.
- M. Strevens, “No understanding without explanation”, in *Studies in History and Philosophy of Science* 44, 2013, pp. 510–515.
- V. Volterra, “Biological associations of two species”, in *Variations and fluctuations in the numbers of coexisting animal species* (1927), reprinted in F.M. Scudo, J.R. Ziegler (eds.), *The Golden Age of Theoretical Ecology: 1923-1940*, Berlin & Heidelberg, Springer, 1978, pp.73–121.
- L. Zagzebski, “Recovering Understanding”, in M. Steup (ed.), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, Oxford, Oxford University Press, 2001, pp. 235–249.

