Micro-level model explanation and counterfactual constraint

Samuel Schindler
Aarhus University

Abstract

Relationships of counterfactual dependence have played a major role in recent debates of explanation and understanding in the philosophy of science. Usually, counterfactual dependencies have been viewed as the explanantia of explanation, i.e., the things providing explanation and understanding. Sometimes, however, counterfactual dependencies are themselves the targets of explanations in science. These kinds of explanations are the focus of this paper. I argue that "micro-level model explanations" explain the particular form of the empirical regularity underlying a counterfactual dependency by representing it as a *physical necessity* on the basis of postulated microscopic entities. By doing so, micro-level models rule out possible forms the regularity (and the associated counterfactual) *could* have taken. Micro-model explanations, in other words, *constrain* empirical regularities and their associated counterfactual dependencies. I introduce and illustrate micro-level model explanations in detail, contrast them to other accounts of explanation, and consider potential problems.

Keywords: explanation, counterfactual dependence, micro-model, modality, necessity

1 Introduction

Philosophers of science often view counterfactual dependence relations, such as "had x been different, y would have been the case", as central to explanations in science – not least since Woodward (2003). More specifically, counterfactual dependence relations are viewed as part and parcel of the explanantia of explanations, i.e., the things doing the explanatory work. But counterfactual dependencies do not always serve as explanantia; sometimes they can serve as explananda, i.e., the targets of explanations. Such explanations are the focus of this paper.

In Woodward's tremendously popular counterfactual account of causal explanation, a causal explanation involves the correct identification of a counterfactual dependence between two variables X and Y, such that a possible intervention on the 'cause' variable X would change the ‘effect’ variable Y (Woodward 2003, 2018).[1] A causal explanation, for Woodward, then provides understanding by allowing us to see how the explanandum (Y) depends on the explanans (X) and, equivalently, under what changes of the explanans the explanandum would have changed. Woodward also speaks of locating the explanandum “in a space of alternative possibilities” and as allowing us to answer so-called “what-if-things-had-been-different” questions, or simply w-questions (Woodward 2003, 191).

Although, for Woodward, causal explanations do not need to explicitly cite generalizations between the cause and the effect variable, “in many, perhaps most scientific contexts, generalizations (laws, causal generalizations etc.), explicitly describing how the explanandum-phenomenon depends on conditions cited in the explanans*, are naturally regarded as part of explanations that the various sciences provide*” (Woodward 2018, 120, changed emphasis). Woodward requires that the generalizations underlying counterfactual dependence relations be *invariant* under interventions, so that the generalization involving X and Y *continues* to hold under some range of interventions on X.

A recent trend in the explanation literature has been to not restrict counterfactual dependencies to those picking out causal relations. For example, in Reutlinger’s “monist” counterfactual theory of explanation is supposed to capture both causal and non-causal explanation (Reutlinger 2016). On Reutlinger’s account, an explanation consists of two parts: a statement describing the explanandum phenomenon and an explanans consisting of at least one generalization (and ‘auxiliary statements’ describing boundary conditions and the like). For the explanans to explain the explanandum, three conditions must be satisfied, one of which is the “dependency condition”, according to which the generalizations of the explanans must support a counterfactual. Several other recent accounts of explanation, likewise, identify “non-causal” counterfactual dependencies as central to an explanation’s explanans (Bokulich 2011, Saatsi and Pexton 2013, and the contributions to the volume edited by Reutlinger and Saatsi 2018).

I do not doubt that counterfactual dependencies supported by generalizations do important explanatory work in science. But science does more than use such generalizations for the purpose of explanation: science sometimes also makes those self-same generalizations the target of its explanations. One way in which science explains generalizations is by invoking models that postulate micro-entities; these models explain why the relationships described by generalizations *must* take the form that they do take (given the model assumptions related to the postulated micro-entities) by *representing contingent generalizations as necessities*. By the same token, such “micro-models” explain the particular form of the counterfactual supported by the relevant generalization. In other words, micro-model explanations *constrain* counterfactual dependencies.

---

[1] Woodward lists a number of conditions that need to be satisfied for a permissible intervention on the cause variable. For more details see Woodward (2003, 98).

I begin the paper (Section 2) by discussing an example that has figured prominently in many discussions in the explanation literature, namely the ideal gas law.[2] The ideal gas law lends itself easily to a counterfactual analysis – and indeed it has been used by Woodward as an illustration of his counterfactual account. But in physics the ideal gas law is often treated not so much as an explanans but rather as an explanandum explained by the kinetic theory of gases, and later statistical mechanics. Woodward has something to say about this, albeit not much positive.

In Section 3 I introduce my account of *micro-level model (MLM) explanations* and illustrate it on the basis of the explanation the kinetic theory of gases provides for the ideal gas law. With three further examples, I demonstrate the wider applicability of my account. In Section 4 I discuss how the MLM model compares to other accounts of explanation with regards to (i) regularity explanation, (ii) the kinds of understanding MLM explanations provide, (iii) in what ways MLM explanations may be said to be reductive (or not), (iv) the role of deduction, and (v) the notion of mechanisms. In Section 5 I further spell out the notion of representation of generalizations (or *regularities*, as I will prefer to call them) as necessities, which is central to my account. I will also address several worries about the notion. In Section 6 I turn to the fact that MLM explanations postulate idealized entities when explaining their target. Allowing for such models to provide genuine explanation raises the specter of explanatory anarchism, which I argue can be avoided. Section 7 concludes the paper.

2 The ideal gas law: explanans or explanandum?

The ideal gas law (IGL) combines the empirically discovered Boyle's law ($P \propto \frac{1}{V}$) and Gay-Lussac's law ($P \propto T$) and has the following form: $PV = nRT$, where P=pressure, V=the volume of the gas container, T=temperature, R=the ideal gas constant, and n=the amount of substance of gas in moles. IGL is briefly discussed by Woodward in his book (Woodward 2003). For example, Woodward at one point in chapter 5 states that, on the basis of IGL, one can explain why, when the temperature of a gas is increased, the pressure of the gas will increase as well, because one can manipulate the pressure and answer a range of w-questions (p. 221). Woodward admits, though, that "this information [provided by IGL], in itself, tells me nothing about *why* an increase in temperature produces an increase in pressure" (p. 222; added emphasis). IGL is explained by statistical mechanics, Woodward continues, but

> Statistical mechanics does not explain in virtue of doing something different in kind from this [what IGL does], but instead simply provides information that allows us to answer what, in some respects, is a wider, more detailed range of w-questions; hence the sense that in some respects, it provides deeper explanations. (For more on this subject and the "in some respects" qualification, see section 5.12.). (Woodward 2003, 223)

Although Woodward does not spell out what he might mean with statistical mechanics providing "deeper" explanations than IGL, in his general view explanations are deeper (than other

---

[2] See e.g., Friedman (1974), Salmon (1984), Woodward (2003), de Regt and Dieks (2005), Elgin (2007), Strevens (2008), Doyle et al. (2019), Rice (2019), Sullivan and Khalifa (2019).

explanations) if they provide answers to a wider range of w-questions (see Woodward 2003, chapter 5). And explanations provide answers to a wider range of w-questions, in turn, if the relationships in question are invariant under a wider range of interventions (than other relationships might be). The van der Waals equation $\left[P+\frac{a}{V^2}\right][V-b]=RT$, for example, is invariant under a wider range of changes than IGL: it incorporates a parameter for the diameter of the relevant gas molecules (b) and a parameter for the attractive forces between the molecules (a) and is accurate for higher degrees of pressure and temperature than IGL. Woodward therefore considers the van der Waals equation to provide a deeper explanation than IGL (Woodward 2003, 260).

Of course, the van der Waals equation is not statistical mechanics, but if Woodward is right, then what is true of the van der Waals equation should also be true of statistical mechanics (and perhaps even to a greater degree). Yet, in the aforementioned quote, Woodward also flags a qualification regarding statistical mechanics providing a deeper explanation than IGL. This qualification he spells out in terms of what he calls the (hypothetical) “microscopic strategy”, which would consist in trying to explain changes in e.g., the value of the pressure variable P upon a change in the volume variable V by considering the initial and final positions of the molecules in the container and the sum of their individual transfer of energy (Woodward 2003, 231-232). The microscopic strategy would try to explain the new value of P by “aggregating the energy and momentum transferred by each molecule to the walls of the container”.

Woodward deems the microscopic strategy unsatisfactory because it fails to answer the relevant w-questions: it doesn’t tell us determinately what value P would have had, had the initial microstate been different. That is so because any macroscopic state (such as a particular value of P) is compatible with basically any initial state of molecules and its evolution (with different molecule trajectories). Woodward points out that the micro strategy would be “impossible” to carry out but concludes that even if it were to be attempted, the microscopic strategy “would fail to provide the explanation of the macroscopic behavior of the gas we are looking for”, because it “omits information that is crucial to an explanation of pressure” (namely counterfactual information) (Woodward 2003, 232). Since IGL, in contrast, does identify situations under which things would have been different, it does a better job than the microscopic strategy at explaining the change in pressure.

From all of this, one might get the sense that “going micro” does not provide much understanding over and above the understanding one already possesses by virtue of IGL.[3] This, however, would clearly be at odds with how scientists usually view the situation. As a look in any textbook introducing the kinetic theory of gases will reveal, IGL is introduced as a (contingent)

---

[3] Another example of his is the macro-explanation of the rise in the price of oranges in terms of a shortage of supply, as opposed to the (fictional) micro-explanation of the same phenomenon in terms of the fundamental physics of “human behavior involved in oranges selling at price P” (233). See Bradley (2020) for a discussion of the claim that “higher level” explanations are generally better because they omit details.

empirical law that then receives an explanation by the kinetic theory of gases.[4] For example, Ebbing and Gammon write in their standard introductory textbook *General Chemistry*: "One of the most important features of kinetic theory is its explanation of the ideal gas law" (Ebbing and Gammon 2017). But what is this "extra" understanding that micro-models such as the kinetic theory of gases provide? In what follows, I will argue that this "extra" understanding consists of micro-models providing explanations of why regularities obtain in the first place and, accordingly, why they support certain counterfactuals, and not others.

3 Micro-level model explanation and counterfactual constraint

In this section I will first provide a definition of micro-level model explanations, or MLM explanations, for short (3.1). I will then illustrate the MLM account with the explanation of the ideal gas law by KT (3.2) and several other examples (3.3).

3.1 MLM explanations defined

MLM explanations consist of an explanandum regularity (supporting counterfactuals) and an explanans consisting of a "micro-level model" that makes idealized assumptions about the microscopic physical system underlying the explanandum regularity and explains the regularity (and its associated counterfactual) by representing it as a physical necessity. More specifically:

> <u>Explanandum</u>: MLM explanations target empirically discovered regularities between macroscopic variables supporting counterfactuals.[5] More precisely, MLM explanations target the *contingent form* of a regularity. One of the simpler (and probably most common) forms of a regularity between two macroscopic variables is a relation of *proportionality* or *inverse proportionality*: when X and Y are proportional, Y increases when X increases; when X and Y are inversely proportional, Y decreases when X increases, or vice versa.[6]

The explanans of an MLM explanation has the following three components:

1. <u>Representation of co-variation</u>: a micro-level model postulates microscopic variables that represent the macroscopic variables in the explanandum regularity so that the two kinds of variables *co-vary*, that is, a change in the values of a macroscopic variable corresponds to a (positive or negative) change in a microscopic variable in the model, and vice versa.

2. <u>Representation of physical necessity:</u> derivations from the model's postulates allow the model to represent the *contingent form* of the explanandum regularity (holding between

---

[4] See also the textbooks by Levine (1995), McQuarrie and Simon (1997) and Holton and Brush (2001), which includes an extensive historical introduction. See Strevens (2008) for perhaps the most extensive philosophical discussion of KT's explanation of IGL (but see also Section 4.1).

[5] Explanandum regularities of MLM explanations can also be probability distributions. See Section 5.3

[6] Another form could be a quadratic relation so that low or high values in X result in low values of Y, but middle range value X results in high values of Y.

macroscopic variables) as *a physical necessity* (holding between microscopic variables).[7] Through this representation, the micro-level model *constrains* the form of the explanandum regularity and its associated counterfactual: the model tells us why, given the micro-level model, the explanandum regularity and its associated counterfactual must have the form that they do have *and not some other form* (e.g., proportionality instead of inverse proportionality).

3. *ww*-questions: component 1 and 2 allow the model to answer *why-would-things-have-been-different questions*, or simply *ww*-questions (pronounce: "quadruple-u questions"): the model provides *reasons for why* changes in the macroscopic variables occurring in the explanandum must happen *by necessity*.

Let us consider a short abstract example. Say there is a regularity involving macroscopic variables X and Y with proportional form. Why does the relation between X and Y have the form that it does have? A micro-model would answer this question by postulating microscopic variables X* and Y* and by relating X to X* and Y to Y*, so that both X-X* and Y-Y* co-vary, and by representing the *contingent* proportionality between X and Y as a necessity by means of X* and Y*. The model can then answer a *ww*-question of the sort of "*why* would Y have increased (rather than decreased) upon increases in X?" by pointing to the microscopic relations of the model, where a decrease in Y* upon an increase in X* *would not have been possible*. This is basically the scheme that the explanation of IGL by the kinetic theory of gases (KT), as we'll now see in detail.

3.2 How KT explains IGL

As already noted, KT's explanandum, IGL, relates macroscopic variables of pressure, temperature, and volume of a gas and supports counterfactuals such as: "had the temperature changed (and the volume stayed fixed), the pressure would have changed" or "had the volume changed (and the temperature remained fixed), the pressure would have changed". More specifically, IGL has a certain form: T and P are related proportionally and V and P *inversely* proportionally (with the appropriate variables fixed). IGL is an *empirical* "law", i.e., it was discovered on the basis of experiments (by Boyle, Gay-Lussac, and others), and as such, it is entirely contingent: in a different possible world, it could have taken a different form, such as $P = nRTV$. In such a world, when the volume of a gas is decreased (and T held fixed), the pressure of the gas would decrease (rather than increase, as in IGL).[8] In other words, contrary to actual fact, V and P could have been related

---

[7] I speak of physical necessities because the necessities are meant to hold about the physical world. They are different from metaphysical necessities or mathematical necessities (such as the necessity involved when Mother fails to distribute 23 strawberries among her three children; see Lange (2017)). See also Section 5 for more on necessities.

[8] In standard usage, the attribute "contingent" applies to objects (such as IGL) which depend on how our actual world is like / happens to be like. That is how I intend to use this attribute also in this paper. Contingency should not be confused with arbitrariness, though. For example, the famous "all gold spheres are less than 1 mile in diameter" is contingent, but also arbitrary and is therefore generally considered not to be a law of nature. Quite in contrast to "all uranium spheres are less than 1 mile in diameter", which supports counterfactuals (such as "had there been spheres of 1 mile in diameter, they would just have blown

proportionally. Let me refer to this hypothetical law as the *other worldly gas law*, or just OWGL. A question one can then ask is why, in our world, are pressure, temperature, and volume in our world related by IGL, rather than by OWGL? I will argue that KT gives an answer precisely to this question. More specifically, I will argue that KT explains IGL by representing it as a physical necessity.

In what follows I will first briefly summarize how IGL is explained in standard textbooks (3.2.1). I will then outline how my account accommodates this example (3.2.2).

3.2.1 A textbook derivation

Let us first of all note that KT is a model in every sense of the word: it makes a number of idealizing assumptions, most notably that molecules are perfectly elastic and that there is therefore no energy loss when the molecules collide with the sides of gas container. Also, it is assumed (very unrealistically) that the molecules don’t interact with each other, either by collision or by repulsion forces. Quantum mechanical effects are ignored and the theory of choice is classical mechanics.[9]

In standard textbooks, the explanation of IGL is presented in the form of a derivation from model assumptions, as follows.[10] Let l be the distance between two opposite container walls and v_x = the molecule velocity along the x-axis. The time elapsed between a molecule traveling from one side of the container (A) to an opposing side of the container (B) and back is $^{2l}/_{v_x}$ and the reciprocal quantity ($^{v_x}/_{2l}$) then is the number of collisions (with B) per time interval. The change of momentum for each molecule is then determined by multiplying the number of collisions per time interval (i.e., $^{v_x}/_{2l}$) with the momentum change per collision, which is $2mv_x$, because the velocity of the molecule changes from positive to negative after a molecule’s collision with the wall, thus $\Delta mv_x = \, mv_x - (-mv_x) = 2mv_x$. The change of momentum is therefore $m{v_x}^2/l$ per molecule per time interval. The total momentum change per time interval for *all* molecules (N) in the container with average speed $\langle {v_x}^2 \rangle$ is simply $N\frac{m\langle {v_x}^2 \rangle}{l}$. Since by Newton’s second law the rate of change of momentum is equal to a force, the average net force exerted by the molecules on a container wall per time interval is: $\langle F \rangle = N\frac{m\langle {v_x}^2 \rangle}{l}$. Since pressure is equivalent to force per area, the pressure exerted by the molecules perpendicularly on the side of the container B is $\bar{F}$ divided by this area of the container (here: l^2), namely $P = N\frac{m\langle {v_x}^2 \rangle}{l^3}$. Since l^3 is just the volume of a cubical container, so we can write $P = \frac{Nm\langle {v_x}^2 \rangle}{V}$. Assuming that there is no preferred direction of a

---

up immediately”). It is of course those latter kinds of regularities that I’m interested in here as targets for micro model explanations.

[9] See Strevens (2008) for a detailed discussion of the idealizing assumptions of KT. KT is called a theory, rather than a model, only for historical reasons.

[10] In my derivation I follow mostly Holton and Brush (2001). For very similar derivations see the textbooks by Levine (1995), McQuarrie and Simon (1997), which are both cited by Strevens (2008). In contrast, Doyle et al. (2019) use a partition function to derive the IGL. A partition function gives the sum over all energy states of the system, and as such is quite different from a derivation in KT that is based on the (classical) dynamics of the individual molecules (although information about molecules of the gas, such as their number, also enters the calculation of the partition function).

molecule's path in the three spatial dimensions, the average velocity of a molecule in the x-axis direction $\langle {v_x}^2 \rangle$ is then just a third of the average velocity of each molecule ($\frac{1}{3}\langle {v_x}^2 \rangle$). Plugging this in our equation of P, we get $PV = \frac{1}{3}Nm\langle {v_x}^2 \rangle$. Since the total kinetic energy of translation KE_{trans} of all molecules is $N \times \frac{1}{2}m\langle {v_x}^2 \rangle$, this gives us $PV = \frac{2}{3}KE_{trans}$, also known as the *gas pressure equation*. This equation is considered extremely significant: it relates (some of) the macroscopic variables of the explanandum to the microscopic variables of the model.

In a separate derivation it can be shown that $T = \frac{2}{3k_B}KE_{trans}$ (whereby k_B = the Boltzmann constant), i.e., the temperature of a gas is proportional to the mean kinetic energy of translation per molecule of the gas. This expression is equivalent to $\frac{1}{2}m\langle v^2 \rangle = \frac{3}{2}k_B T$. If we multiply this expression with Avogadro's number N_A (=number of molecules per mole, which is a constant) we get $\frac{1}{2}N_A m\langle v^2 \rangle = \frac{3}{2}RT$. Since $N_A m$ = the molar mass M of a gas, we get $\frac{1}{3}M\langle v^2 \rangle = RT$, which together with the gas pressure equation then completes KT's representation of the ideal gas law as a necessity.

3.2.2 Representation, co-variation, physical necessity, and *ww*-questions

In the derivation of IGL from KT we find all the elements of MLM account present. First, we can see that the macroscopic variables of IGL, namely T, V, and P are being represented as microscopic variables in the model, namely molecular speed, intermolecular space (or gas density), and molecule-wall collisions, respectively. We can also see that these macroscopic and microscopic variables co-vary (positively and negatively): (i) when the temperature of the gas increases (decreases), KT represents this as the molecular speed increasing (decreasing), (ii) when the volume decreases (increases), KT represents this as the gas density increasing (decreasing), (iii) when the pressure increases (decreases), KT represents this as the frequency of molecule-wall collisions increasing (decreasing).

Second, the derivations from the model assumptions concerning highly idealized molecules (as detailed in the last section) allow KT to represent IGL as a physical necessity: conditional on the physical assumptions of the model, IGL holds by necessity. KT thus tells us why IGL rather than OWGL holds, and why e.g., P and V are related inversely proportionally (rather than proportionally).

Third, by virtue of representing IGL's macroscopic variables in terms of microscopic variables and by virtue of representing their relations as physical necessities, KT allows us to answer *ww*-questions. For example, we can answer the question of "*why* would P have increased, rather than decreased, had the volume been reduced (and the temperature held fixed)?" The answer KT gives is: if the volume of the gas had been reduced, this would co-vary with an increase in gas density, and less space for the same number of molecules in motion (whilst keeping their speed fixed) *must* result in an increase in the frequency of molecule-wall collisions, which in turn represent an increase in gas pressure. This can clearly be seen from the gas pressure equation $PV = \frac{2}{3}KE_{trans}$, which we derived in the last section: when the volume V of the container is reduced, pressure P will have to increase by virtue of the magnitude of the translational kinetic energy of the molecules being inversely proportional to the magnitude of the volume (the smaller V, the bigger $\frac{2}{3}KE_{trans}$ and therefore P). It is therefore *not possible*, by virtue of KT's representation of the macroscopic variables in IGL, for the gas volume to be reduced and P to decrease. In contrast, on IGL alone V and P are related inversely proportionally only contingently, i.e., V and P could also have been related differently than how they are related by IGL (see e.g., the OWGL). By accounting for IGL in this way, KT helps us go beyond the answers we obtain by asking (Woodwardian) w-questions about IGL: KT provides *reasons for why* the variables of IGL are related in the way they are related. KT thus takes as its explanatory target an object (namely IGL) that other philosophers have often taken to be the primary source of explanation.

3.3 Further examples

In this section I want to demonstrate the applicability of the MLM account on a basis of three further examples (in this order): the Bohr model of the atom and the explanation of the spectral lines of hydrogen, Dalton's atomism and the explanation of the laws of constant proportions, and Mendel's explanation of hybridization experiments. For each I will show that the essential conditions of the MLM account are satisfied.

3.3.1 The Bohr model

The Bohr model of the atom (proposed in 1913) was devised to explain the empirically discovered spectral line patterns of hydrogen, which were summarized in the so-called Rydberg formula $\frac{1}{\lambda} = R(\frac{1}{{n_1}^2} - \frac{1}{{n_2}^2})$, which describes the spectral emission and absorption lines of hydrogen. Starting with the work by a schoolteacher by the name of Balmer, formulas for line series were discovered by trial and error and n did not have any deeper meaning.[11] In order to explain the phenomena summarized by the Rydberg formula, Bohr postulated that, inside the atom, electrons revolve around the nucleus on stable orbits with fixed energy levels. Whenever electrons would "jump" from one orbit to another, the atom would emit or absorb light (depending on whether the "jump" was from a higher to a lower energy level or vice versa). On the basis of the Coulomb's law and several other assumptions,[12] Bohr was then able to derive a model-based representation of the Rydberg formula. The constant R could now be specified in terms of the concepts of the model, such as electron mass, charge, the Coulomb constant, and π (for circular orbits). The quantum numbers n_1 and n_2, which in the Rydberg formula had no deeper physical meaning, could now be interpreted as the "order" of electron orbits (starting from orbit closest to the nucleus), with n_1 as the 'ground state' of a series (n_1=2 for the Balmer series) and n_2 (=3, 4, 5 … for the Balmer series) as the permitted, quantized higher energy states. The Bohr model allows us to calculate the energy levels and the radii of the orbits associated with these levels.

What is the analysis that the MLM account would suggest of the explanatory power of the Bohr model? First of all, it is apparent that the Rydberg formula supports certain counterfactuals,

---

[11] In the Rydberg formula, n_1 determines the kind of spectral line series, e.g., the Balmer series for n_1=2.

[12] For example, Bohr assumed – quite boldly – that the frequency of emitted light (which is inversely proportional to its wavelength) is equal to the average of the frequencies of the electron on its orbit before and after the "jump".

such as "had n_2 been changed, the wavelength λ of emitted / absorbed light would have changed". As such, the formula thus would allow us to answer Woodwardian w-questions. Although this is indeed useful for scientists (following Balmer's lead, other spectral series were discovered and described), it is not fully satisfactory. Scientists want to know *why* empirically discovered regularities hold. The Bohr model provides such an answer. First, the Bohr model postulates microscopic entities, namely electron orbits, which represent variables (n_1 and n_2) of the explanandum regularity, whereby the microscopic and 'observed' variables co-vary: the bigger (smaller) n, the higher (smaller) the energy level and the bigger (smaller) the radius of an electron orbit. Now the *form* of the empirical regularity captured by the Rydberg formula, namely the emission or absorption of light when n_2 is changed, the Bohr model explains by the quantum jumps of the electron between its orbits: the emitted or absorbed wavelengths *has got* to be quantized by *physical necessity*, because the electron orbits are discrete; it's *not possible*, on the Bohr model, for electrons to move continuously between orbits, and there therefore cannot be any continuous emission or absorption of light. The Bohr model thus allows us to answer *ww*-questions of the form "*why would* the emitted or absorbed wavelengths have changed non-continuously, had the n values been different".

The Bohr model happens to be one of the examples used by the prominent account of model explanation according to which fictional models explain their targets by answering Woodwardian w-questions based on counterfactual dependencies (Bokulich 2011). For example, "Bohr's model is able to correctly answer a number of 'what-if-things-had-been-different questions', such as how the spectrum would change if the orbits were elliptical rather than circular …" (Bokulich 2011, 43). Several issues arise from making fictional entities the objects of counterfactual dependencies (Schindler 2014, Nguyen forthcoming). For example, Nguyen highlights an ambiguity in Bokulich's account between first- and second-order explanations. Roughly, the former in Bokulich's account would amount to fictional models correctly capturing counterfactual relations pertaining to the target system (CRT), e.g., the tides (A) depend on the masses and positions of celestial bodies (B). Second-order explanations, on the other hand, would amount to the fictional models explaining CRTs in the target, i.e., they would accurately represent what it is that A's dependence on B itself depends on (e.g., gravitational forces). Nguyen argues that Bokulich's account most plausibly is read as targeting second-order explanations and, further, that fictional models cannot provide such second-order explanations, because CRTs *cannot* depend on the fictional features of the model (which do not exist). Since my account of explanation does not presuppose any counterfactual dependence between the model and the target, Nguyen's criticism does not apply.[13]

---

[13] One of my own points of criticism of Bokulich's account draws attention to a tension between Bokulich's demand that model fictions be justified by true theories in order for them to count as explanatory on the one hand, and her view that model fictions can provide deeper explanations than true theories on the other hand. If explanatory model fictions must be justified by true theories, however, then the explanations provided by such models ultimately will counterfactually depend on the explanans identified by those true theories (and the model explanation can therefore not be deeper than the explanations provided by those

3.3.2 Dalton's model

Now enter Dalton's atomism (proposed in 1808), which at its heart consisted of the idea that chemical elements decompose into atoms specific to each element and that atoms are indivisible. The Daltonian model explains the laws of definite proportions and the law of multiple proportions. The former law says that chemical elements always combine in the same ratios. For example, oxygen always makes up 8/9 and hydrogen 1/9 of the mass of water. The latter law says that whenever two elements form more than one compound, the ratio of those compounds will be multiples of each other, as carbon monoxide and -dioxide where a certain amount of carbon combines with exactly twice as much oxygen in the second compound as it does in the first. A counterfactual associated with these laws might be e.g., "had the amount of carbon been x′ (rather than x), the amount of oxygen would have been y′ (rather than y)".

Dalton's atomism explains the *form* of the laws and their associated counterfactuals, i.e., it explains not only the fact *that* we observe what is described by the two laws, but also why we do *not* observe any intermediate ratios. E.g., we would not observe that one unit of carbon combines with 1 ½ as much oxygen (rather than twice as much) as it does in CO. Such a combination is ruled out by the assumption of indivisible atoms: it is for this reason that elements combine with each other only in integer multiple proportions.[14] In sum, we can say that Dalton's atomism explains *ww*-questions such as "*why would* the amount of carbon have doubled in a sample of carbon monoxide, had the amount of oxygen doubled (rather than tripled)?" The answer the model gives is that, since atoms are indivisible, there *could not have been* combinations of elements with different proportions.

3.3.3 Mendel's model

Let us consider a final example. In hybridization experiments with pea plants, Mendel made some ground-breaking discoveries in pea crossing experiments in the mid-19th century. When he crossed purebred (recessive) white and (dominant) purple flower plants and self-fertilized their offspring, the second filial generation (F2) would consist of a 3:1 ratio of purple vs. white flowered plants. A counterfactual associated with that law might be: "had this cross of flowers consisted of *only* pure-bread dominant or recessive, a 4:0 ratio of either purple or white flowers would have resulted".

Mendel explained the law and the associated counterfactuals by invoking indivisible unitary genetic 'factors' (as he called them) that come in pairs for each organism (alleles). For each hybridized pair of plants there are exactly four possible combinations of alleles (two per plant).[15]

---

true theories) (Schindler 2014, 1747). This point of criticism strikes me as being rather similar to the second horn of Nguyen's dilemma (Nguyen forthcoming, 3240) – contrary to what Nguyen's remarks in footnote 32 of his paper may suggest. (The first horn of Nguyen's dilemma is the point mentioned above, namely that the target cannot counterfactually depend on fictions.).

[14] In view of the second point of our definition in Section 3.1, we can say more precisely that Dalton's atomism *represents* each chemical element (the macroscopic variable) with its own kind of atoms (the microscopic variable), all of which are indivisible, and which therefore combine with atoms of other elements only in multiple integers.

[15] Again, we may say more precisely that Mendel's model *represents* phenotypic traits (the macroscopic variable) in terms of alleles (the microscopic variable).

Together with the principle of dominance, according to which traits of the dominant genetic 'factors' are always expressed phenotypically, the self-fertilization of purebred recessive and dominant plants from F1 *must* result in a 3:1 ratio in F2. For comparison, a 2:2 ratio in F2 would be impossible on the model. Once again, also Mendel's model answers *ww*-questions such as "*why would* the proportion of red and white flowers in F2 have been 3:1 (rather than e.g., 3.5:2.3), had one crossed red and white plants in F1?". The answer in the model is: because the relevant traits are inherited discretely on dominant and recessive alleles, making it impossible to obtain ratios that are combinatorially incompatible with a set of 2 x 2 alleles.

I take these three examples to sufficiently motivate the applicability of the MLM account.

4 MLM and other accounts

In this section, I will issue several clarifications with regards to how the MLM account compares to other accounts of explanation with regards to (i) the explanation of regularities (4.1), (ii) the kind of understanding MLM occasions (4.2), (iii) how MLM compares to reductive explanations (4.3), (iv) the role of deduction (4.4), and (v) the notion of mechanisms (4.5).

4.1 Regularity explanation

As mentioned in the introduction, invariant generalizations (such as IGL) and the counterfactual structures they support are primary vehicles of explanation in Woodward's influential account (Woodward 2003). Similarly, in the classical Deductive-Nomological (DN) model of explanation, laws together with boundary conditions constitute the explanans, from both of which the explanandum is deduced and thereby explained (Hempel and Oppenheim 1948, Hempel 1965). In contrast, accounts that take such generalizations or laws as objects of explanation are far and between. As Strevens has noted: "the literature on regularity explanation is remarkably thin" and "philosophers have for the most part provided only programmatic comments on the explanation of [regularities or] laws" (Strevens 2008, 209). Instead, the focus of accounts of explanation has been on *events* as objects of explanation.

For example, even though the DN model of explanation was originally meant to apply also to the explanation of regularities, it was soon realized that this version of the DN model would face a problem that is sometimes referred to as the "problem of conjunction" (Hempel and Oppenheim 1948, see also Salmon 1989): from the conjunction of b = Boyle's law (concerning gases) and k = Kepler's laws (concerning the planets) one can logically derive k. Although this example seems to fit the DN model, we do not recognize this derivation as an explanation: it is no explanation to derive k from k plus an irrelevant conjunct (such as b). This problem remained unresolved by proponents of the DN model and they henceforth "restrict[ed] their attention to the explanation of particular facts" (Salmon 1989, 9). The MLM account arguably avoids the problem of conjunction, as the derivation involved in MLM explanations has to be based on a model, which plausibly must obey coherence constraints, ruling out "laws" (or rather relations between microscopic variables) irrelevant to the matter at hand. Also, MLM explanations require the postulation of microscopic variables and their relations representing the explanandum regularity, ruling out a scenario in which a law is invoked to explain itself.[16]

Even Strevens, who identifies regularity explanation as a topic worth addressing, believes that regularity explanation reduces to event explanation. Or as he puts it, causal-mechanistic accounts of explanation, such as his own kairetic account, have the advantage of "obviating the need to say anything special or particular about regularity explanation at all" (Strevens 2008, 228). More specifically, for Strevens, to causally explain e.g., why All Fs are Gs just boils down to causally explaining why F-ness causes G-ness (see also Strevens 2012). Hence, for Strevens, "… the target of a law-explaining model is not the law to be explained", but rather the "property that completes the instantiation of the law", namely G-ness (Strevens 2008, 226). In other words, on Strevens' account we can explain why All Fs are Gs by explaining why (all) instances of F cause instances of G. That is why the explanandum (namely the regularity) and the explanatory target, on Strevens' account, are "entirely different kinds of entity" (Strevens 2008, 226). This can also be seen from Strevens' detailed discussion of a part of the ideal gas law in chapter 8 of his book, namely Boyle's law: the explanatory target, for him, is not the law itself but rather what he calls "Boylean behavior" of gases, i.e., a gas's pressure and its *property* of "its having a magnitude related to the gas's volume by Boyle's law, that is, a magnitude of k/V" (Strevens 2008, 311). In contrast to Strevens' reductive account of regularity explanations, on the MLM account of explanation regularities are explanatory targets sui generis.[17]

4.2 Understanding

On the traditional DN model of explanation, understanding was equated with "nomic expectability" (Salmon 1989). As Hempel wrote, "the [DN] argument shows that, given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected;* and it is in this sense that the explanation enables us to *understand why* the phenomenon occurred" (Hempel 1965, 337, original emphasis). Causal and mechanistic accounts, which came to dominate thinking about explanation in subsequent decades, are intended to provide "how-actually" understanding of how the phenomenon was produced by its underlying causes or mechanisms (Salmon 1984, Craver 2007). Woodward (2003), the no doubt most influential account of explanation in recent years, has made an impressive and lasting case for understanding in scientific explanations involving counterfactual understanding. More recently, there has been

---

[16] Strevens (2008) seeks to solve the puzzle by requiring that explanatory principles describe the aggregation of real causal influences and by ruling out cases as the Kepler-Boyle example as causally irrelevant to the explanation (pp. 275-6). Other attempts are discussed in Salmon (1989). See also Kitcher (1976).

[17] An anonymous referee referred me to the account by Doyle et al. (2019). Doyle et al. introduce the distinction between an *object* and a *base* of understanding, whereby they define the former as the "thing to be understood" and the latter as the "thing to provide understanding". As one of their examples, they consider the explanation of the IGL (an object of understanding) by the so-called partition function in statistical mechanics. However, Doyle et al. do not provide any detail as to how a *philosophical* account of regularity "understanding", let alone *explanation*, might look like. Instead, their primary goal is to defend non-factivism about understanding, that is, the view that "radical departures from the truth are not always barriers to understanding" (345) (to which I myself would subscribe to).

much talk about models providing 'how-possibly' understanding (Grüne-Yanoff 2013, Rohwer and Rice 2013, Reutlinger et al. 2018, Massimi 2019), where models just identify *possible* explanatory factors.

What kind of understanding is provided by MLM explanations? First, although the MLM account shares the view with the DN account that deduction is an important part of scientific explanations, it does not locate understanding in rendering the explanandum (an event) "expected", but rather in representing the explanandum (a regularity) as *necessary*. More specifically, MLM explanations tell us why a regularity must obtain (conditional on the model assumptions) and why, consequently, only certain counterfactuals hold (and not others). The MLM account can therefore be said to provide (conditional) *how-necessarily* understanding. I will have more to say about the modal nature of the MLM account later on (Section 5).

Second, MLM explanations answer *ww*-questions and thereby give us understanding of the *changes* occurring in the variables of the explanandum regularity (that are related contingently) in terms of the *changes* in the microscopic variables (which are related by necessity) postulated by the model. Even though the MLM account differs from Woodward's account with regards to a number of issues, such as (i) identifying the vehicles of explanation and understanding (generalizations on Woodward's account, models on the MLM account), (ii) the kind of modality involves in explanation (counterfactual modality on Woodward's account and physical necessity on the MLM account), and (iii) a commitment to realism (Woodward's account relies on causal relation in the world, whereas the MLM account doesn't), the MLM account shares the view with Woodward that the ability to "relate changes" is important for understanding the explanandum.[18]

4.3 Reductionism

MLM explanations are reductionist in nature. It is an essential part of the reductionist program that the phenomena of interest are explained and understood by "bringing them back to" more basic phenomena, which tend to be at the microscopic scale (e.g., molecules, atoms, fundamental particles) (van Riel and Gulick 2019). Likewise, MLM explanations seek to explain the relationships holding between macroscopic variables, i.e., measurable quantities describing the gross state of a system, *in terms of* microscopic variables, which are taken to *constitute* the macroscopic state of a system. In our example, the relationships holding between the macroscopic variables *pressure* and *temperature* are explained by a model that postulates molecules whose (highly idealized) properties and behaviors are considered to constitute these variables (pressure is constituted by molecule-wall collisions and temperature by molecular speed). And as we saw, the explanation proceeds by representing the macroscopic relationships *in terms of* microscopic

---

[18] Woodward (2003) introduces the notion of "change-relating relationships" between explanans and explanandum which he defines as relating "changes in variables in the explanans to changes in the explanandum variable" (202). Again, for Woodward these change-relating relationships mostly consist of generalizations such as IGL (rather than relationships between a model and a regularity).

entities. This is indeed what one would expect from a reductionist explanation.[19] So there is a sense in which MLM explanations may seem compatible with reductionism. A special aspect of MLM explanation, though, is that the entities postulated by MLM explanations are highly idealized: we know they don't exist in the form postulated by the respective models (e.g., molecules *do* have an extension and they *do* interact). On some accounts of reduction, this would render MLM explanations not actual reductions. The standard Nagelian account of reduction, is one such account.

Nagel's original account of reduction is simple and straightforward: a theory T1 reduces to a theory T2 if the laws of T1 can be deduced from the laws of T2 and some auxiliary assumptions and so-called 'bridge laws' connecting the different kinds of ('heterogeneous') vocabularies of T1 and T2 (Nagel 1961).[20] As such, the Nagelian account might be viewed as compatible with the MLM account of explanation, particularly given the close association of the Nagelian account with the classic DN model of explanation ('if you reduce, then you explain', a slogan might go). There are a couple of rather substantial differences though. First of all, the Nagelian model of reduction is supposed to hold between *theories* (such as the classical example of thermodynamics and statistical mechanics); it is not a model of how theories explain empirical regularities to begin with. But even ignoring this rather basic point, the MLM account of course says much more than what Nagelian reduction is about. For example, the MLM account explains why certain regularities (and their associated counterfactual) *and not others* hold between certain macroscopic variables (such as *P* and *V*) and *why* certain changes in macroscopic variables would have *had to* happen upon changes in other macroscopic variables. Neither 'change-relating relations', nor the representation of contingent explanandum regularities as physical necessities, are part of the Nagelian picture.

Finally, Nagel himself and others came to believe that theory reduction always involves the *correction* of the reduced theory by the reducing theory.[21] But of course, a reducing theory can only correct a reduced theory if the reducing theory itself is true (or closer to the truth in the right respects). Clearly, that is not the case with a highly idealized model such as KT. If anything, KT makes *more* incorrect assumptions about the target system than the empirical regularity that it explains (IGL holds up until certain high ranges of temperature and pressure, and that could be viewed as an assumption of IGL).

---

[19] In virtue of their reductive nature, MLM explanations contrast quite sharply with so-called "minimal model" explanations, as identified by Batterman and Rice (2014). Minimal "minimal models" in science explain their targets (some macro-behavior) on the basis of the model and the target being members of the same universality class. Many of the microscopic details of the systems considered does not matter to the explanations provided by minimal models. For criticisms of Batterman and Rice's account see e.g., Lange (2015). Incidentally, I agree with Lange that there *does* seem to be a sense in which minimal models in their own examples explain their targets by sharing some (abstract) features (contra Batterman and Rice).

[20] Nagel's original account seems to come with all kinds of positivistic presumptions (e.g., a distinction between observational and theoretical language), but it has been argued more recently that the basic idea of Nagelian reduction does not require such baggage (Dizadji-Bahmani et al. 2010).

[21] A reason for that belief is the idea that "often, it is in fact not possible to derive the *exact* laws [of the reduced theory]", as for example in the second law of thermodynamics (Dizadji-Bahmani et al. 2010, 398).

4.4 Deduction

An obvious similarity between the MLM account and the DN model of explanation is that in MLM, too, a logical deduction is involved in the explanation of the explanandum. But this is where the similarities already end. I see three differences between the two accounts.

First, a *workable* version of the DN model of explanation is available only for the explanation of *events*, such as the fall of an apple (see Section 4.1). As mentioned earlier, the explanatory targets of the MLM account, in contrast, are empirical regularities. Second, on the DN model the explanans explains the explanandum simply by virtue of logically deducing the explanandum. On the MLM account, in contrast, the explanandum regularity is *represented* by (co-varying) microscopic variables and relations that are derived from the model assumptions. In KT's explanation of IGL, for example, the relation between gas pressure and temperature (Boyle's law) is explained by representing that relation as a necessary relation between molecule-wall collisions and molecule speed. The representation of the explanandum regularity allows the model to provide answers to *ww-questions*, i.e., why-would-things-have-been-different questions. For example, KT tells us why (as part of IGL) the counterfactual dependence between pressure and volume is such that had the volume been reduced, the pressure would have risen. Answer: because a reduction in intermolecular space *necessarily would have resulted* in an increase in the frequencies of molecule-wall collisions (corresponding to the macroscopic variable pressure*)*. In contrast, the DN model merely aims to provide answers to simple *why* questions (e.g., why did the apple fall from the tree?). Third, on the DN model it is required that the laws be true for them to be explanatory. There is no such requirement on the MLM account. On the contrary, MLM explanations are based on models that idealize and simplify reality. We will return to this issue in Section 6 of this paper. Let us now turn to mechanistic explanations.

4.5 Mechanisms

On mechanistic accounts of explanation, the explanandum phenomenon (which can be a regularity) is explained by reference to a mechanism that "produce[s], underlie[s], or maintain[s] the phenomenon of interest" (Craver and Tabery 2019). Mechanisms have been defined differently by different authors, but mechanisms are often described as consisting of entities, or parts, and their activities. Obviously, MLM explanations also invoke entities and (sometimes) their activities to explain the explanandum. In that sense, MLM account is superficially similar to mechanistic accounts. There are however at least two crucial aspects in which MLM explanations diverge from some of the fundamental assumptions in the mechanism literature.

First, in mechanistic explanations there is no necessity involved: mechanisms explain by showing how the explanandum phenomenon was *actually* brought about. For example, an explanation of the neural action potential involves providing neurobiological detail about the *actual* opening and closing of *actual* ion channels etc., not just how the action potential might possibly be brought about (Craver 2007). On the MLM account, in contrast, the representation of empirical regularities as *necessities* is essential for the understanding obtained with micro-models. This representation of empirical regularities as necessities, in turn, allows MLM explanations to provide answers to *ww*-questions: by virtue of them we know *why* certain counterfactuals hold about the real system in question and others don't, and why if variables in the explanandum were changed, such changes *by necessity* (rather than by contingent fact) had to result in changes in other variables in the explanandum.

It is worth noting that the change-relating understanding associated with the MLM account (Section 4.2) is different from Craver's mechanistic account (Craver 2007). Craver suggests that Woodwardian counterfactuals be used to individuate mechanistic explanations, consisting of properties or activities ϕ of the components of a mechanism and "aggregate" of those properties or activities ψ, which at the same time constitutes the explanandum phenomenon: if we were to manipulate ϕ, ψ would change, and, vice versa, if we were to manipulate ψ, ϕ would change. Craver calls this the 'mutual manipulability' criterion (Craver 2007, 154-159). It is quite apparent that if the entities postulated by a model do not exist (as in micro-level models), then there cannot be a mechanistic explanation according to this criterion; non-existent properties or activities cannot be manipulated. But even looking past that and acknowledging a similarity between Craver's mutual manipulability and the idea of co-variation of variables in the MLM account (if variables co-vary, then they also support counterfactuals), the difference is still that on the MLM account the form of the explained regularity is represented as necessary, that is, the possibility is ruled out that the regularity could have taken any other form. That is a (logically) stronger view than the mechanists' view, which describes how the form the explanandum *happened to take* was actually brought about (rather than how *it must be*).

Second, and related to the first point, several proponents of the mechanistic view are realists about mechanisms: they are committed to the entities and activities constituting mechanisms being real (Machamer et al. 2000, Craver 2007, Illari and Williamson 2011, Craver and Tabery 2019). The MLM account, in contrast, is more liberal: often the entities postulated by micro-models are knowingly unrealistic, as in all of the examples discussed above. That is, contra KT, there are intermolecular collisions; contra the Bohr model, there are no electron orbits in atoms; contra Dalton, atoms are surely divisible; and contra Mendel, gene expression is often much more complicated than producing a phenotype from two pairs of alleles.[22]

Some accounts in the mechanistic literature are described as 'epistemic' and have been distinguished from 'ontic' accounts. The former are often associated with the work of Bechtel and the latter with the work of Craver. Ontic accounts stress that mechanisms are part of the actual "causal structure of the world" rather than just representations (Craver 2007). Ontic accounts of mechanistic explanation are thus firmly realist. Epistemic accounts, in comparison, view explanation as "fundamentally an epistemic activity performed by scientists" (Bechtel 2008). Illari (2013) points out that neither view is "instrumentalist" and describes the difference as one that boils down to whether the description or "exhibition" of a mechanism entails ontic or merely epistemic commitments. Be that as it may, the "core consensus" of proponents of both kinds of accounts, according to Illari, lies in the acceptance of the view that mechanisms *produce* the phenomena they explain (Illari 2013, 239). However, a production relation wouldn't be a good

---

[22] See the literature on reduction of Mendelian genetics to molecular biology (see Waters 2013).

characterization of the kinds of relations that hold between the model and the explanatory target in several of the micro-model explanations that I've discussed here. For example, gases as described by KT do not actually produce thermal phenomena (because gases are much different in reality), the Bohr model does not actually produce spectral phenomena (because atoms are not small solar systems), and eye color is not actually produced by pairs of dominant and recessive alleles (rather, no less than 15 genes have been associated with the production of eye color).

Even though the main mechanistic accounts of explanation are realist accounts, Colombo et al. (2015) have argued that talk about mechanisms in science may well be interpreted instrumentally. So instead of viewing mechanisms as entities that require our full realist commitments (in every detail), they can instead be viewed as models that (i) are supported by the evidence (to one degree or another), (ii) are coherent with background knowledge, and (iii) facilitate instrumental success with regards to the explanandum phenomenon. Still, there is "one core aspect of mechanistic explanation" that even Colombo et al. do not question in their antirealist rendering, namely, again, that mechanisms *produce* the phenomena they explain (Colombo et al. 2015, 196). But, again, a production relation is problematic as a characterization of the relation that holds between micro-model explanations and the phenomena they explain. It is perhaps possible in principle to construe the production relation in instrumentalist terms, but I do think that giving up on a realist reading of it would radically change the mechanistic project beyond recognition.[23]

In sum, the MLM account is substantially different from other superficially similar accounts of explanation. In particular, (i) the assumptions made by MLM explanations are often not true, (ii) MLM explanations do not provide how-actually explanations, but rather represent empirical regularities as necessities, and thereby allow us to answer ww-questions, and (iii) the relation between the micro-level model and the explanandum cannot be described as a (mechanistic) production relation. The first and the third point concern the issue of idealization, which is of course prevalent in models. As already mentioned, I will have more to say about this in Section 6. In what follows I will first seek to shed more light on the role of necessity in MLM explanations.

5 Necessity in the MLM account

According to the MLM account, necessity plays a crucial role in explaining regularities and their associated counterfactuals. In what follows, I will further elucidate this role.

5.1 What kind of necessity?

What kind of necessity is involved in the explanations provided by the MLM account? The short answer is: physical necessity that is conditional on the truth of the model assumption. But let me provide a longer answer.

In modal logic, the necessity operator (often symbolized with the 'box' symbol '$\Box$') indicates that the statement it ranges over holds in all possible worlds. That is obviously not the kind of

---

[23] I am not *in principle* opposed to using the label of "mechanistic explanation" for MLM explanations: so long as the basic elements of the features identified by my account are respected (see Section 3.1), this would then merely be a terminological matter.

necessity to be found in MLM explanations, as there clearly are worlds in which the model is false (for example, *our* world!). The necessity of MLM explanations, rather, is intended as a physical necessity. Physical necessities are more restrictive than logical possibilities, or for that matter, metaphysical possibilities.[24] Physical necessities are usually defined in relation to the laws of nature: what is physically necessary follows from the laws of nature of the actual world (Vaidya 2017). For example, it is physically not possible for anything to move faster than the speed of light, because one of the fundamental laws of physics rules that out. In other words, it is physically necessary that any moving object moves with speeds not exceeding the speed of light (if it is not possible that not-*p*, this is logically equivalent to that it is necessary that *p*).

Again, the world in which the assumptions of MLM explanations are true is not the actual world. So how can we understand the physical necessities of MLM explanations? In a sense it's quite simple: if the world in which the model assumptions hold were true, the relations described by the model would be physically necessary. That is why the physical necessities of MLM explanations are *conditional* physical necessities. For example, if the assumptions of the Bohr model were true, and electrons were to traverse only fixed orbits around the nucleus, it would not be possible for atoms to emit or absorb energy continuously; it would be physically necessary for atoms to emit or absorb energy discretely. Or if the assumptions of the kinetic theory were true, and molecules perfectly elastic point particles, it would not be possible for gases to expand upon their temperature increasing; it would be physically necessary that this would happen. Deductive logic helps us establish these necessities on the basis of the model assumptions.

Because necessity is so central to the MLM account, it may be referred to as a modal account of explanation. Yet, the MLM account is quite different from another recent modal account of explanation, namely Marc Lange's account of "explanation by constraint" (Lange 2017). Explanations by constraint, according to Lange, work "by identifying certain constraints to which the world must conform" (Lange 2017, xvi). For example, mathematical facts constrain the way in which a mother will fail to distribute 23 strawberries evenly among her three children (by mathematical necessity) and conservation laws constrain the kinds of physical forces that are possible (by physical necessity), namely only those that conserve energy. These latter kinds of constraints concerning physical laws, in particular, are presumably supposed to exist 'in the world'. In contrast, the physical necessities of micro-level models are only conditional on the truth of model assumptions.

5.2 Necessities and contingencies

According to the MLM account, we understand why a certain regularity and its associated counterfactual holds, because we, in a sense, "replace" the contingency of the regularity with the necessity that derives from the model. One may object, however, that on certain accounts, regularities, or "the laws of nature", are necessities themselves, conceived of as relations between universals (e.g. Armstrong 1983). But if laws are themselves necessities – the thought continues –

---

[24] Logical possibility is believed to be the most encompassing modality there is. Metaphysical modality is believed to more restricted, but the details are disputed. See Vaidya (2017).

then a central tenet of my account seems to be undermined, namely that an MLM explanation consists of representing a contingent regularity as a necessity. In other words, MLM explanations could then not increase our understanding modally, because the explained law *already* is necessary.

I have three comments about this objection. First, the nomic necessity view of lawhood does not sit particularly well with the fact that many laws of nature are not exceptionless universal generalizations of the form "all Fs are Gs" (as supposed by the view) but rather ceteris paribus laws that hold only under certain conditions and only within certain ranges (Reutlinger et al. 2017). For example, as we mentioned already, the ideal gas law is correct only up until a certain temperature and pressure (above which the van der Waals equation is more accurate). So either the ideal gas law is not a law (which would be absurd), or the nomic necessity view of lawhood is too strict for describing the regularities that often are the target of MLM explanations.

Second, many have objected that the notion of "non-logical or contingent" necessities seems to be rather mysterious (and perhaps self-contradictory); proponents of the nomic necessitation view have failed to provide arguments for the existence of such relations (Lewis 1983). Third, even if such necessitation relations existed, they would still be (metaphysically) *contingent*. That is, on the nomic necessitation view of lawhood it would still be contingent that, in our world, a reduction in gas volume necessitates an increase in gas pressure. That must be so, because otherwise we wouldn't need science to discover the laws of nature. Hence, even on accounts that view the laws of nature as relations of necessitation (not mine!), micro-level models could still be appealed to in order to explain why the (somewhat mysterious) *contingent* necessitation relation has the form that it does have in our world.[25]

In sum, I think the nomic necessitation view of laws of nature is either so problematic that the laws of nature better not be viewed as necessities, or the nomic necessitation view actually entails that laws of nature are contingent anyway (despite the necessitation relation). In either case, the core idea of the MLM account remains intact.

5.3 Necessities and events

Salmon once raised an objection against the "modal conception" of explanation, i.e., the conception of explanation that "says that a good explanation shows that what did happen had to happen" (Salmon 1998, 321). Salmon conceived of the modal conception in terms of a relation between a physically necessary law of nature and an event-to-be-explained, which made it sound quite akin to the DN model. He stressed, though, that in contrast to the DN model, on the modal conception the explanandum is not explained by deriving it from a law, but rather by showing that the explanandum event is "physically necessary relative to the explanatory facts" (Salmon 1984, 111). At any rate, the objection Salmon raised against this version of the modal account is in principle

---

[25] The same point holds for antireductionist views such as Maudlin's, which stipulate the necessity of the laws of nature as a primitive in their metaphysics (Maudlin 2007).

also relevant to modern modal accounts, such as Lange's (2017) and also for my MLM account. So let us consider it in some detail here.

Salmon's complaint was that the modal account cannot accommodate statistical or even truly indeterministic explanations that can be found throughout modern science; Mendelian genetics and quantum mechanics are cases in point. For example, quantum mechanics does not predict (or explain) any particular measurement of a quantum system; it only predicts certain probabilities of measurement outcomes. Likewise, Mendelian genetics does not predict (or explain) any the color of any particular pea flower, but only certain probability distributions (of how likely it is that the pea flower in a particular plant will blossom white). Thus, modal accounts that suppose that *events* are explained by reference to a physically necessary law (as a version of the DN model that Salmon considers does) indeed don't look very plausible in the face of such examples.

An obvious rejoinder to this kind of criticism is that some scientific theories or models explain probability distributions, rather than any particular event. Salmon (1998, 323) actually pondered this reply himself but dismissed it by arguing that sometimes science really is interested in the explanation of single events, or at least a small number of events (e.g., Rutherford wanted to explain why a small number of alpha particles was deflected at large angles when directed at gold foil). But even if we grant that sometimes scientists seek to explain single events or (as in Salmon's examples) a small number or a "restricted" number of events, it is still true that the explananda in many explanations in science are regularities and not any (particular, single) event. Whenever that's the case, it would be strange to demand from an account of explanation that it accommodates the explanation of single events – even though the scientific explanation itself does not.[26] For illustration, let us re-consider Mendel's model of inheritance (which, incidentally, is also Salmon's example against modal accounts) in a bit more detail.

Mendel's model predicts that the probability distribution of flower color in the aforementioned example of the hybridization of pure-bred pea plants will be 3:1 (see Section 3.3). Thus, any given flower in F2 has a 1 in 4 *probability* of being white. This distribution is indeed necessitated by Mendel's model: provided the 'experiments' are conducted conscientiously, and provided the sample is large enough to support statistical inferences to the population mean, there cannot be any other distribution in F2. Again, this is so because there are exactly four possible combinations of alleles (two per plant), and with one type of allele being dominant and the other type of allele (namely the allele for white blossom) being recessive, the result *has got to be* 3:1. A distribution of 2:2, for example, is not possible on the model. But because the model necessitates *probability* distributions, the model does not necessitate the color of any *particular* plant, or what a statistically underpowered, small number of breeding experiments will yield. These two issues are worth treating in more detail.

---

[26] There is a sense in which the explanation of regularities automatically entails the explanation of single events, since regularities are based on (or even grounded in) events. I take this sense of explaining events when one explains regularities to be trivially true. Salmon certainly had something else in mind when raising his challenge to the modal account.

First, one might very well be interested in why a particular plant has the color that it has. Presumably one would look for an explanation to this question by studying the DNA and the gene expression in that particular plant. Mendelian genetics does not provide an answer to this question. This of course does not mean that it's not a legitimate question to ask, but it's simply a different a question to the one Mendelian genetics seeks to answer. Second, a limited number of breeding experiments may or may not yield trait distributions which are not precisely the ones predicted by Mendelian genetics. For example, a limited set of experiments on the hybridization of pure-bred pea plants may not yield the 3:1 distribution but something slightly else, for example, 2.99:1.01. Would this undermine Mendelian genetics or the view (that follows from my MLM account) that the ratio in F2 must be 3:1? The answer is no. The distributions predicted by Mendelian genetics, by virtue of being *probability* distributions, are *ideal limits*. With a limited number of experiments these ideal limits may not actually be achieved.[27] But that's just a property of statistics and the law of large numbers. There may also be all sorts of impediments in the actual experiments that cause divergences from the ideal limit.[28] As important as it is to hold these impediments and confounding factors under control, they are not under the purview of the explanation provided by Mendelian genetics.

All in all, modal accounts, and the MLM account in particular, are not undermined when not accommodating the explanation of particular events.[29] And they shouldn't be, because the scientific explanations that they seek to capture are not undermined in this way either.

6 Idealization and explanatory demarcation: beyond the dichotomy

The MLM account of explanation has no truth requirement: the postulated entities (and their activities) are not required to be correct. An obvious worry is that the MLM account results in explanatory anarchism where "any explanation goes". While I reject explanatory anarchism, I am happy to embrace explanatory liberalism. In fact, I believe it's unavoidable: the truth requirement is not compatible with actual scientific practices regarding explanation.

---

[27] It bears some irony that Mendel himself has been suspected of bias because he presented numbers from his experiments which seemed too good to be true (Franklin et al. 2008).

[28] Note that the Duhem thesis seems also relevant to this point: theories are never applied to the world without a number of auxiliary assumptions about the experimental setup, the experimental apparatus, background effects, etc. Cartwright (1983) also speaks of a "prepared (i.e. idealized) description" of the phenomena which theories explain (rather than the nitty gritty of the actual world).

[29] The analogous considerations – in principle – apply to quantum mechanics, Salmon's other alleged counterexample to the modal account: quantum mechanics predicts probabilities of *possible* measurement outcomes with necessity, but it doesn't predict any particular measurement outcomes. I would be prepared to argue that it is (also) by virtue of necessitating these probabilities that quantum mechanics possesses explanatory power. Of course, quantum mechanics is a special case, since the veritable measurement problem poses its very own difficulties for *any* account of explanation. See e.g., Salmon (1998, 325). It's not clear how even Woodward's more liberal account of causation could accommodate quantum mechanics, and in particular, the Einstein-Podolsky-Rosen paradox (de Regt 2004).

Traditionally, accounts of scientific explanation used to require that the explanans be true for it to explain the explanandum phenomenon (Hempel 1965). As we mentioned earlier (Section 4.5), mechanistic accounts of explanation, too, require this. Yet the traditional dichotomy of "either truth, or no explanation (and understanding)" has been challenged in the past decade by whole trove of works (Elgin 2004, 2007, Strevens 2008, Batterman 2009, Bokulich 2011, Rohwer and Rice 2013, de Regt 2015, Potochnik 2017, Reutlinger et al. 2018, Doyle et al. 2019, Rice 2019). These works stress that explanations and understanding in science often are based on idealizations and literally false assumptions. A possible way of accommodating these practices lies in giving up on a *strict* truth requirement, without giving up on the truth requirement altogether.[30]

Approximate truth is a much-used notion in the scientific realism debate. Attempts to formalize the notion have turned out difficult (see e.g. Chakravartty 2017), but the basic idea is clear enough: no realist believes that our best current scientific theories capture reality fully accurately, but only to some extent. Something along those lines might be said of scientific models: only those micro-level models are explanatory that capture the approximate truth about the postulated entities underlying micro-level explanations.[31] Such a view would certainly help to accommodate some of the examples discussed here: KT captures the approximate truth about gases consisting of molecules in motion (although it misrepresents many other features of molecules); Dalton's theory of atoms captures the approximate truth about chemical reactions consisting of the combination of atoms (although it misrepresents atoms as being indivisible); Mendelian genetics captures the approximate truth about inheritance being based on genes (although it oversimplifies inheritance). Yet, the view that model explanations explain by virtue of capturing the approximate truth regarding their target faces a worry, which would be analogous to a worry raised in the scientific realism debate.

In the scientific realism debate, it is often debated exactly which parts of a theory were actually needed to generate empirically correct predictions (Vickers 2013). Likewise, a general worry about attributing a model's explanatory power in its truth might be that a model's idealizations are actually essential for its explanatory power. In fact, that's precisely what has been argued for a number of idealizations in physics and biology (Batterman 2009, Kennedy 2012, Rice 2015). For example, sometimes gases are falsely modeled as continuous fluids in order to account for 'shocks' in gases (i.e., areas of high molecular density within a gas created by external pressure). Without this idealization, Batterman (2009) contends, there wouldn't be any explanation of the target phenomenon. Similarly, Rice (2015) argues that false assumptions in "optimality

---

[30] There are other solutions to this dilemma. Strevens (2008) suggests that idealizations in reality just flag those causal factors that don't make a difference to the explanandum. Bokulich (2011, 2012) argues that model explanations are justified through "translation keys" with true theories. See Schindler (2014) and Nguyen (forthcoming) for criticisms of the role of these translation keys in Bokulich's account. See also Section 3.3.1.

[31] Thanks to Ioannis Votsis for suggesting this to me. See also Pincock (forthcoming) for a 'veritist' proposal along those lines.

explanations" in biology, such as the assumption that populations are infinite and that organisms mate randomly, cannot be removed from the model "without consequently eliminating the explanation being offered" (601). Moreover, Rice (2019) argues more generally that models typically cannot be decomposed into accurate and inaccurate parts, but are instead to be viewed as "holistically distorted representations", because "the contributions made by the purportedly accurate parts of the model only make their contributions within an idealized mathematical modelling framework that pervasively distorts those feature" (191). One of Rice's examples so happens to be the kinetic theory of gases and the ideal gas law. Thus, if idealizations are indeed indispensable to model explanations, then the truth-requirement of explanations cannot be retained.

There is yet another reason for giving up on the truth-requirement for explanations: we would be in a much better position to make sense of why certain models in the past were adopted for their perceived ability to provide explanation and understanding. Consider the notorious caloric theory of heat (CT), in which heat is conceived of as a *substance*, called caloric. Caloric particles are mutually repulsive but attract (and are attracted by) matter particles (Figure 1).

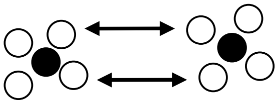

Figure 1: Caloric particles (white) repel each other but are attracted by matter particles (black).

Most interestingly for our purposes, CT can represent IGL and the associated counterfactuals as a necessity on the basis of its postulated entities (just like KT does; see Section 3). Consider for example this counterfactual supported by IGL: had the gas's volume been reduced, the pressure would have risen. According to CT, the counterfactual has this rather than another form (e.g., a pressure reduction upon the reduction of the gas's volume) because the mutual repulsion of caloric particles will be stronger the closer the particles are to each other (which will be the case when the volume is reduced). Or consider another counterfactual dependence: had the temperature of the gas increased (while holding fixed the volume of the gas), the pressure would have increased. On CT, the counterfactual has this rather than another form (e.g., a pressure increase upon a temperature increase), because an increase in the amount of caloric in the container (associated with an increase in temperature) necessarily leads to an increase in the net amount of repulsion between the caloric particles, which, in turn, is associated with an increase in pressure of gases on CT. So it seems that CT, despite its patent falsity, can explain IGL in a very similar way as KT can: it too answers *ww*-questions.

One could of course simply take a conservative stance towards CT and insist that since caloric does not exist, CT did not explain at all. But, again, that would make the acceptance of CT by the contemporary scientific community at the time as an explanatory model a mystery (see e.g. Chang 2003). Alternatively, one could try to render the change from CT to KT one of structural continuity, despite a radical change in the postulated entities (Votsis and Schurz 2012). I am skeptical of such a rendering.[32]

In the spirit of explanatory liberalism, and against the dichotomy of "either truth or no explanation", I want to suggest that we simply accept that a model like CT is explanatory. At the same time there are very good grounds for thinking that KT is a *much better* explanatory model than CT. Those grounds have to do with the *theoretical virtues* of KT. For instance, KT has much wider *explanatory scope*, as it explains not only IGL, but also the properties of substances in different states, heat transfer and conduction of gases. Similarly, KT has made many successful predictions (such as the specific heat ratios of gases (de Regt 1996)) and has overall provided a very *fertile* research program (Clark 1976). In contrast, CT soon ran into problems that it could not solve, such as the apparently indefinite production of heat in the boring of cannons (as famously pointed out by Count Rumford) in contradiction with CT's central tenet that heat is a conserved substance (i.e., it cannot be destroyed or just disappear) or the mere fact that – as a substance – it appeared imponderable. Thus, by virtue of its wider explanatory scope, greater fertility, and inherent contradictions, KT *overall* turned out to be a much better explanatory model than CT. It would therefore be irrational to insist on using CT instead of KT to explain IGL.

Explanatory scope and fertility are of course only two out of a number of theoretical virtues on the basis of which scientists can assess their theories and models (Kuhn 1977, Schindler 2018). Simplicity and mathematical tractability are also important considerations, ruling out more fantastical 'models' with a much more demanding metaphysics than what's required by e.g., postulating particles and their interactions.

What emerges from these considerations is that scientific explanation is not a matter of "truth or no explanation", but rather more of a continuum with plenty of differentiation between good and terrible (or no) explanations. Theoretical virtues help us choose the theories that overall are the better theories; realists and antirealists can continue to argue about whether these virtues give us reason to think that those best theories are likely to be true, or whether we are better served epistemologically to suspend judgments about this matter.

7 Conclusion

Counterfactual dependence is widely seen as fundamental to scientific explanation. However in this paper I argued that there are kinds of explanation in science that take as their targets the regularities that support counterfactuals. These kinds of explanations therefore explain what counterfactual accounts of explanation take for granted, namely the particular form of the counterfactual and its underlying regularity. I call these explanations micro-level model explanations, or MLM explanations, for short, because they achieve their explanatory goals by postulating microscopic variables. On the basis of these microscopic variables, I argued, the model

---

[32] The structure retained in this example seems a purely empirical structure we know by virtue of IGL. But structural realism requires structural continuity at the *theoretical* level. See Worrall (1989).

represents the contingent target regularities as necessities, thereby constrains the relevant counterfactuals, and provides answers to *why-would-things-have-been-different questions*. I argued that MLM explanations are not captured by some of the leading accounts of explanation.

References

Armstrong, David Malet. 1983. *What is a Law of Nature?* Cambridge: Cambridge University Press.

Batterman, Robert W. 2009. Idealization and modeling. *Synthese*, **169** (3): 427-446.

Batterman, Robert W and Collin C Rice. 2014. Minimal model explanations. *Philosophy of Science*, **81** (3): 349-376.

Bechtel, William P. 2008. *Mental mechanisms: Philosophical perspectives on the sciences of cognition and the brain*. Oxford: Routledge.

Bokulich, Alisa. 2011. How scientific models can explain. *Synthese*, **180** (1): 33-45.

———. 2012. Distinguishing Explanatory from Nonexplanatory Fictions. *Philosophy of Science*, **79** (5): 725-737.

Bradley, Darren. 2020. Should Explanations Omit the Details? *The British Journal for the Philosophy of Science*, **71** (3): 827-853.

Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford: Oxford University Press.

Chakravartty, Anjan. 2017. Scientific Realism. *Standford Encyclopedia of Philosophy*, edited by Edward N. Zalta, <https://plato.stanford.edu/archives/sum2017/entries/scientific-realism/>.

Chang, Hasok. 2003. Preservative realism and its discontents: Revisiting caloric. *Philosophy of science*, **70** (5): 902-912.

Clark, Peter. 1976. Atomism versus theromodynamics. In *Method and appraisal in the physical sciences: The critical background to modern science, 1800-1905*, Colin Howson (ed.), Cambridge: Cambridge University Press.

Colombo, Matteo, Stephan Hartmann, and Robert Van Iersel. 2015. Models, mechanisms, and coherence. *The British Journal for the Philosophy of Science*, **66** (1): 181-212.

Craver, Carl F. 2007. *Explaining the brain : mechanisms and the mosaic unity of neuroscience*. Oxford: Oxford University Press.

Craver, Carl F. and James Tabery. 2019. Mechanisms in Science. *The Stanford Encyclopedia of Philosophy (Summer 2019 Edition)*, edited by Edward N. Zalta, <https://plato.stanford.edu/archives/sum2019/entries/science-mechanisms/>.

de Regt, Henk W. 1996. Philosophy and the Kinetic Theory of Gases. *The British Journal for the Philosophy of Science*, **47** (1): 31-62.

———. 2004. Review of James Woodward, making things happen. *Notre Dame Philosophical Review*:

———. 2015. Scientific understanding: truth or dare? *Synthese*, **192** (12): 3781-3797.

de Regt, Henk W. and D Dieks. 2005. A contextual approach to scientific understanding. *Synthese*, **144** (1): 137-170.

Dizadji-Bahmani, Foad, Roman Frigg, and Stephan Hartmann. 2010. Who's Afraid of Nagelian Reduction? *Erkenntnis*, **73** (3): 393-412.

Doyle, Y, S Egan, N Graham, et al. 2019. Non-factive understanding: A statement and a defense. *Journal of General Philosophy of Science*, **50** (3): 345-365.

Ebbing, Darrell and Steven D Gammon. 2017. *General chemistry*. Eleventh Edition ed. Boston: Cengage Learning.

Elgin, Catherine Z. 2004. True enough. *Philosophical issues*, **14** (1): 113-131.

———. 2007. Understanding and the facts. *Philosophical Studies*, **132** (1): 33-42.

Franklin, A., A.W.F. Edwards, D.J. Fairbanks, et al. 2008. *Ending the Mendel-Fisher Controversy*. Pittsburgh: University of Pittsburgh Press.

Friedman, Michael. 1974. Explanation and scientific understanding. *The Journal of Philosophy*, **71** (1): 5-19.

Grüne-Yanoff, Till. 2013. Appraising models nonrepresentationally. *Philosophy of Science*, **80** (5): 850-861.

Hempel, Carl G. 1965. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.

Hempel, CG and P Oppenheim. 1948. Studies in the Logic of Explanation. *Philosophy of Science*, **15** (2): 135-175.

Holton, Gerald James and Stephen G Brush. 2001. *Physics, the human adventure: From Copernicus to Einstein and beyond*. New Brunswick: Rutgers University Press.

Illari, Phyllis. 2013. Mechanistic explanation: Integrating the ontic and epistemic. *Erkenntnis*, **78** (2): 237-255.

Illari, Phyllis and J Williamson. 2011. Mechanisms are real and local. In *Causality in the sciences*, Phyllis Illari, F Russo and J Williamson (eds.), Oxford: Oxford University Press, 818 - 844.

Kennedy, Ashley Graham. 2012. A non representationalist view of model explanation. *Studies in History and Philosophy of Science Part A*, **43** (2): 326-332.

Kitcher, Philip. 1976. Explanation, Conjunction, and Unification. *The Journal of Philosophy*, **73** (8): 207-212.

Kuhn, Thomas S. 1977. Objectivity, Value Judgment, and Theory Choice. In *The Essential Tension*, Chicago: University of Chicago Press, 320-333.

Lange, Marc. 2015. On "minimal model explanations": A reply to Batterman and Rice. *Philosophy of Science*, **82** (2): 292-305.

———. 2017. *Because without cause: Non-causal explanations in science and mathematics*: Oxford University Press.

Levine, Ira N. 1995. *Physical Chemistry*. Fourth Edition ed. New York: McGraw-Hill.

Lewis, David. 1983. New work for a theory of universals. *Australasian Journal of Philosophy*, **61** (4): 343-377.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. Thinking about mechanisms. *Philosophy of Science*, **67** (1): 1-25.

Massimi, Michela. 2019. Two kinds of exploratory models. *Philosophy of Science*, **86** (5): 869-881.

Maudlin, Tim. 2007. *The metaphysics within physics*: Oxford University Press on Demand.

McQuarrie, Donald A. and John D. Simon. 1997. *Physical Chemistry. A Molecular Approach*. Sausalito, CA: University Science Books.

Nagel, Ernest. 1961. *The structure of science; problems in the logic of scientific explanation*. New York: Harcourt.

Nguyen, James. forthcoming. Do fictions explain? *Synthese*:

Pincock, Christopher. forthcoming. A Defense of Truth as a Necessary Condition on Scientific Explanation. *Erkenntnis*:

Potochnik, Angela. 2017. *Idealization and the Aims of Science*. Chicago: University of Chicago Press.

Reutlinger, Alexander. 2016. Is there a monist theory of causal and noncausal explanations? The counterfactual theory of scientific explanation. *Philosophy of Science*, **83** (5): 733-745.

Reutlinger, Alexander, Dominik Hangleiter, and Stephan Hartmann. 2018. Understanding (With) Toy Models. *British Journal for the Philosophy of Science*, **69** (4): 1069–1099.

Reutlinger, Alexander and Juha Saatsi. 2018. *Explanation Beyond Causation: Philosophical Perspectives on Non-causal Explanations*: Oxford University Press.

Reutlinger, Alexander, Gerhard Schurz, and Andreas Hüttemann. 2017. Ceteris paribus laws. *Stanford encyclopedia of philosophy*, edited by Edward N. Zatla, <https://plato.stanford.edu/archives/spr2017/entries/ceteris-paribus/>.

Rice, Collin. 2015. Moving beyond causes: Optimality models and scientific explanation. *Noûs*, **49** (3): 589-615.

———. 2019. Models Don't Decompose That Way: A Holistic View of Idealized Models. *The British Journal for the Philosophy of Science*, **70** (1): 179-208.

Rohwer, Yasha and Collin Rice. 2013. Hypothetical pattern idealization and explanatory models. *Philosophy of Science*, **80** (3): 334-355.

Saatsi, Juha and Mark Pexton. 2013. Reassessing Woodward's account of explanation: Regularities, counterfactuals, and noncausal explanations. *Philosophy of Science*, **80** (5): 613-624.

Salmon, Wesley. 1984. *Scientific Explanation and Causal Structure of the World*. Princeton: Princeton University Press.

———. 1998. *Causality and explanation*. New York: Oxford University Press.

Salmon, Wesley 1989. Four Decades of Explanation. *Scientific Explanation: Minnesota Studies in the Philosophy of Science*, **13**: 3–219.

Schindler, Samuel. 2014. Explanatory fictions—for real? *Synthese*, **191** (8): 1741-1755.

———. 2018. *Theoretical Virtues in Science: Uncovering Reality Through Theory*. Cambridge: Cambridge University Press.

Strevens, Michael. 2008. *Depth: an account of scientific explanation*. Cambridge, Mass.: Harvard University Press.

———. 2012. Précis of Depth. *Philosophy and Phenomenological Research*, **84** (2): 447-460.

Sullivan, Emily and Kareem Khalifa. 2019. Idealizations and Understanding: Much Ado About Nothing? *Australasian Journal of Philosophy*, **97** (4): 673-689.

Vaidya, Anand. 2017. The Epistemology of Modality. *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, <https://plato.stanford.edu/archives/win2017/entries/modality-epistemology/>.

van Riel, Raphael and Robert Van Gulick. 2019. Scientific Reduction. *The Stanford Encyclopedia of Philosophy (Spring 2019 Edition)*, edited by Edward N. Zalta, <https://plato.stanford.edu/archives/spr2019/entries/scientific-reduction/>.

Vickers, Peter. 2013. A Confrontation of Convergent Realism. *Philosophy of Science*, **80** (2): 189-211.

Votsis, Ioannis and Gerhard Schurz. 2012. A frame-theoretic analysis of two rival conceptions of heat. *Studies in History and Philosophy of Science Part A*, **43** (1): 105-114.

Waters, Ken. 2013. Molecular Genetics. *The Stanford Encyclopedia of Philosophy (Fall 2013 Edition)*, edited by Edward N. Zalta, <https://plato.stanford.edu/archives/fall2013/entries/molecular-genetics/>.

Woodward, James. 2003. *Making things happen: a theory of causal explanation*. Oxford: Oxford University Press.

———. 2018. Some Varieties of Non-Causal Explanation. In *Explanation Beyond Causation: Philosophical Perspectives on Non-Causal Explanations*, Alexander Reutlinger and Juha Saatsi (eds.), Oxford: Oxford University Press.

Worrall, John. 1989. Structural Realism: The Best of Both Worlds? *Dialectica*, **43** (1-2): 99-124.

Acknowledgments:

I thank the two anonymous of this journal for their detailed comments, which helped to improve this paper. Thanks also to Kareem Khalifa, Daniel Kostić, and Insa Lawler for reading very carefully a distant relative of this paper and for making me re-think. Previous versions of this paper were presented at Workshop on Understanding from Models at the IHPST in Paris (2017) and at the research colloquium at Simon Fraser University (2019). I thank the audiences at both talks for their feedback.