Regulative Idealization: A Kantian Approach to Idealized Models

Lorenzo Spagnesi

Abstract: Scientific models typically contain idealizations, or assumptions that are known not to be true. Philosophers have long questioned the nature of idealizations: Are they heuristic tools that will be abandoned? Or rather fictional representations of reality? And how can we reconcile them with realism about knowledge of nature? Immanuel Kant developed an account of scientific investigation that can inspire a new approach to the contemporary debate. Kant argued that scientific investigation is possible only if guided by ideal assumptions—what he calls "regulative ideas". These ideas are not true of objects of nature, and yet they are not heuristic tools or fictional representations. They are necessary rules governing the construction and assessment of scientific explanations. In this paper, I suggest that some idealizations can be interpreted as having necessary regulative value and as being compatible with scientific realism. I first analyze the puzzle of the nature of idealization and present the main approaches to this topic in the literature. Second, I reconsider the puzzle vis-à-vis a restricted, Kantian definition of idealization and a novel characterization of the relation between idealization and truth. Finally, I discuss in detail an example of idealization (the Hardy-Weinberg equilibrium) along the suggested Kantian lines.

Keywords: idealization, models, Kant, regulative ideas, realism, scientific explanation

1. Introduction

Idealizations, or assumptions that are known not to be true, are ubiquitous in the sciences. A common idealization in biology is that populations of organisms are infinitely large. In economics, scientists often assume that humans behave as perfectly rational agents (*homo economicus*). A frequent idealization in physics (famously employed by Galilei in *The Two New Sciences* published in 1638) is the assumption of frictionless planes. More specifically, idealization seems to have a close connection with scientific modeling. Scientific models, as partial and incomplete representations of phenomena, typically contain idealizations. Frictionless planes, infinite populations, or the *homo economicus* are all idealized models or ideal assumptions of more complex models.[1]

In this paper, I will ask some fundamental questions about idealizations: are they heuristic tools that will eventually be abandoned? Or rather fictional representations of reality? Are there different kinds of idealization? And how can we reconcile their use with realism about knowledge of nature? I will try to answer these questions from a Kantian perspective that hopefully will shed some new light on long-standing puzzles. I suggest that Kant's doctrine of regulative ideas is a precious resource for addressing the contemporary problems of understanding the role of idealization and its compatibility with realism. Although Kant could not have been aware of the multifarious uses of idealizations that can be found in contemporary science, this is not a long shot. There is an interesting and neglected continuity that connects Kant to the contemporary debate in the philosophy of science. The debate about idealizations goes back to the writings of Hans Vaihinger in the early years of the 20th century (Vaihinger 1911). Vaihinger was an eminent Kant scholar and elaborated a fictionalist view of scientific

[1] It may be argued that *all* models involve some form of idealization (see, e.g., Bokulich 2009). In this paper, however, I will remain neutral on this issue.

models based on his interpretation of Kant's regulative ideas. Although Vaihinger deeply influenced the subsequent debate (Fine 1993, Suárez 2009; Bokulich 2016) and there have been recent appeals to Kantian terminology in the literature (e.g. Weisberg 2007),[2] the Kantian origin of the debate has so far not been explored. This paper aims to amend this lacuna. Philosophers of science have been looking at Kant as a source of inspiration for many topics: laws of nature, monism, pluralism, etc. My paper will show that Kant's account of scientific investigation can be a precious resource even for debates about idealizations.[3]

An important influence for this paper is an insightful distinction made by Margaret Morrison in her book *Reconstructing Reality: Models, Mathematics, and Simulations* (2015). She distinguishes between 'abstraction' and 'idealization' in science. Abstraction is "a process whereby we describe phenomena in ways that cannot possibly be realized in the physical world" (Morrison 2015, 20). Idealization instead "involves a process of approximation whereby the system can become less idealized by adding correction factors" (ibid.). She then emphasizes how, while idealization (so defined) is primarily used to ease calculation, abstraction is (at least in some cases) "a necessary condition for explaining and hence understanding the phenomena in question" (ibid., 31). Morrison's terminology is peculiar and intentionally poles apart from standard accounts.[4] But while one may find this terminology uncharacteristic (I take abstraction so defined to be indeed a form of idealization since it involves an ideal description of phenomena), there is a crucial insight that the distinction makes perspicuous and that I want to further explore: some non-realistic, idealized representations are not just calculation tools but necessary ingredients for explanation.

Despite terminological peculiarities, I suggest that Morrison's distinction has a Kantian flavour and that investigating its Kantian background can help us complement her views. More specifically, I will show how Kant's accounts of regulative ideas provides a fruitful springboard to think the implications of Morrison's distinction and make it relevant to the very debate on idealization. I suggest that at least some idealizations can be interpreted as having necessary regulative value and as being in principle compatible with scientific realism. I first analyze the puzzle of the nature of idealization and present the main approaches to this topic in the literature. Second, I reconsider the puzzle vis-à-vis a restricted, Kantian definition of idealization and a novel characterization of the relation between idealization and understanding. Finally, I discuss in detail an example of idealization (the Hardy-Weinberg equilibrium) along the suggested Kantian lines.

2. The puzzle of idealization

Let's start from a standard characterization of idealizations in science (taken from Potochnik 2017):

[2] Weisberg has appealed to "representational ideals" to characterize various kinds of idealizations in an influential paper (Weisberg 2007). On Weisberg's account, some of these ideals (completeness and generality) have "regulative function" (ibid., 649, 654).

[3] I leave the details of this historical connection between Kant and the contemporary discussion to another paper. Here I will focus on some Kantian insights that are promising for the current debate.

[4] Morrison changes the usual referents of the debate to differentiate her account from traditional ones that, on her view, are unable to explain the necessary role of mathematics in explanation (Morrison 2015, 21). In chapter 3 of her book, Morrison provides a detailed account of how mathematics provides genuine physical information. More specifically, she shows how renormalization group methods give us a physical understanding of phase transitions by encoding the relation between microscopic and macroscopic features of empirical systems (ibid., 78–9). My focus in this paper is however not on mathematics as such (although idealizations are often but not necessarily described in mathematical terms), but on how idealizing provides knowledge of phenomena.

> Idealizations are assumptions made without regard for whether they are true and often with full knowledge that they are false. (Potochnik 2017, ix)

One may decide to restrict the domain of idealizations to 'misrepresentations' of empirical phenomena, i.e. false representations only. No matter whether we opt for a more restricted or more liberal version of the characterization, the class of idealizations seems not easily reconcilable with a realist understanding of knowledge of nature, roughly the view that:

> Scientific theoretical conceptions of entities and processes are approximately but literally true. (See e.g. Psillos 1999)

For how can ideal assumptions made without regard to their truth-value be relevant to theoretical conceptions that are meant to be approximately true of nature? This is what I shall call 'the puzzle of idealization'—a puzzle that plays a central role in the heated debate regarding the nature and role of idealizations in science. Note that while the above characterization of realism is 'veritist' in that it assigns a special place to truth in science, it does not imply that truth is the only aim of science. Popper (1963) already objected that, if that were the case, scientists would just attempt to maximize the number of true propositions by formulating extremely probable, if not tautological hypotheses. I maintain, however, that some truth-related criterion (for example, approximate truth or verisimilitude) is a necessary condition of scientific realism, semantically (e.g. Niiniluoto 2017) or epistemically conceived (e.g. Bird 2007).[5]

There is no easy way to solve the puzzle of idealization. Not only does it seem that idealizations are not easily compatible with (or relevant to) truth—they seem to respond to a completely different strategy. In some cases, idealizations are preferred even when more realistic assumptions are available, e.g. when economists use the assumption of *homo economicus* (or better, one of its versions depending on the school of economic thought) to model their theories (as emphasized by Potochnik 2017). If that's correct, it seems that the quest for truth cannot *motivate* idealizations in science at all. As a result, many interpreters either undermine the relevance of idealizations to science or they abandon realism (or veritism) altogether. Either way is costly for (a.) idealizations do seem to be ubiquitous in the sciences; and (b.) they do seem to contribute to our knowledge and understanding of nature. Here, I isolate two major approaches to idealization (roughly following Shech 2018):[6] the heuristic and the fictionalist view, broadly understood. (This is not to deny that others have attempted different strategies—I will discuss some of these attempts below.)

The heuristic view regards idealizations as mere calculation devices that will disappear with the advancement of science. According to this account, idealizations are, as it were, like Wittgenstein's ladders that we throw away once we have climbed up them. Scholars such as

---

[5] Thanks to an anonymous reviewer for urging clarification on this point.

[6] For Shech (2018) the debate around idealizations (more specifically, infinite idealizations in physics) is a contrast between "essentialists" and "dispensabilists": essentialists claim that idealizations are fundamental to scientific accounts of physical systems, whereas dispensabilists think that idealizations can be ultimately eliminated from mature science. Scheck's dispensabilism roughly corresponds to my heuristic view, whereas his essentialism resembles my fictionalism, i.e. the view that idealizations are necessary or essential falsehoods. Note, however, that an idealization could play an essential role without being *used* as a fiction.

Vaihinger (1911), Nowak (1980), Wimsatt (1987), Strevens (2016, 2019), Sullivan & Khalifa (2019) are proponents of this view since it allows to keep a solid notion of truth separated from the employment of false representations. In fact, this view has been recently associated with 'factivism' about explanation: the idea that explanations must be grounded in true, real facts). While I agree that some idealizations fit into this category, the heuristic view seems unable to account for those cases in which idealizing cannot be simply eliminated to obtain a more accurate description. As noted by Morrison, some idealized (or, to use her term, abstracted) representations (e.g., infinite population or thermodynamic limit) seem to offer a kind of understanding that cannot be simply replaced by more accurate explanations. These idealizations are *required* to explain physical phenomena (see Morrison 2015, 27–8). If that's correct, this view is not easily reconcilable with the indispensable function that at least some idealizations seem to play.

The fictionalist view takes idealizations to be indispensable for science—they are ineliminable, yet felicitous falsehoods that advance our understanding in a way similar to other forms of representations in art and literature (Frigg 2010; Elgin 2017; Potochnik 2017). The key move shared by the proponents of this view is that idealizations advance our understanding of phenomena not just despite their falsity but *because* of their falsity. There are many ways one can argue that falsity is necessarily required in science: from real-world complexity to limited epistemic capacities. As a result, however, these interpreters need to reconfigure the epistemic aims of science according to non-realist lines—most commonly, around the idea that understanding is the ultimate epistemic aim of science at the expense of truth (see, e.g., Potochnik 2017). The fictionalist account justifies idealization as a necessary, ubiquitous practice while making idealization epistemically worthwhile. However, this position comes with costs, too. One may be skeptical about the capability of such accounts to deliver a robust notion of scientific knowledge of nature. For these accounts, idealizations are nothing but fictions. If, on the fictionalist view, they are meant to play a key explanatory role, it is at least unclear how scientists can avoid reification of some sort (i.e., mistakenly taking idealizations as referring to 'real' entities, properties, or processes; see section 6 below).

3. The puzzle reconsidered: Kantian idealization

As we saw in the previous paragraph, the two main approaches to idealization seem to leave the puzzle of idealization unsolved. Idealizations are either ultimately dispensable from, as it were, 'mature science' or scientific investigation itself is reconsidered as a non-veritist and non-realist epistemic quest.[7] It is possible, however, to make some headway on this issue if we look more closely at how we defined the problem—more specifically, at our characterization of idealization. Note that the definition of 'idealization' we used is extremely broad. It includes all kinds of misrepresentations and unrealistic assumptions a scientist may use in her investigation—from the intentional misrepresentations contained in Maxwell's vortex-idle wheel model of the ether to the assumption of perfectly rational agent in economics or infinite populations in genetics. But these idealizations are very different from each other. Consider, for example, the case of Maxwell's vortex-idle wheel model of the ether. The ideal nature of this imaginary model played a crucial "exploratory role" in Maxwell's discovery of the equations for the electromagnetic field in the *Treatise on Electricity and Magnetism* (1873).[8] The model itself was, however, abandoned and replaced by the more realistic explanation codified in Maxwell's equations. The same cannot be said of assumptions such as perfect

[7] For an attempt to detach realism from truth, see Potochnik (forthcoming).

[8] As carefully argued in Massimi (2019). For Maxwell, this model of the ether was a "collection of imaginary properties" (Maxwell 1861/2).

rationality in economic behaviour or infinite populations—it is true that they can be replaced by more realistic assumptions (imperfect rationality, finite populations) but that generally implies a loss in explanatory power. As suggested by Morrison, there is simply no Hardy-Weinberg equilibrium, for example, without the assumptions of infinite population and random mating.[9] These assumptions, I suggest, are not simply false but *cannot be true* of nature—[10]they are introduced precisely because of their ideality and are, at least in some sense, 'irreplaceable'. As such, they are difficult to reconcile with the heuristic view about idealizations.

It therefore seems advisable to introduce the following distinction (among other possible ones),[11] namely the distinction between (1) idealizations that are false but can be true of nature from (2) idealizations that are false and simply cannot be true. The first class of idealizations is broad: it includes all idealizations that are potentially true of phenomena and can therefore be corrected by either adding correction factors (as in the case of the Galileian model of frictionless motion) or by correcting core aspects of the model itself (as in the case of Maxwell's model of the ether). The second class of idealizations cannot be either de-idealized nor corrected, and typically contain parameters set to zero or infinity that have no instantiation in nature. The puzzle of idealization should then be reconsidered vis-à-vis such a distinction. In the case of potentially true, yet false assumptions (1), I believe the puzzle is indeed not soluble given the premises. Keeping potentially-true-yet-false representations of phenomena instead of realistic representations would simply go against the realist's agenda. There might be all sort of ways to reconcile these idealizations with realism but, I suspect, they would involve either regarding this kind of idealization as an (ultimately dispensable) tool towards truer explanations or loosening the truth requirement of scientific knowledge (perhaps by emphasizing the cognitive limits of human agents).[12] Neither way would show that idealizations *as such* are necessary and compatible with the scientific quest for truth.[13]

In the case of idealizations that cannot be true (2), the situation may appear even worse. For these assumptions are not even meant to be replaced by new, more accurate ones. The only available solution seems to discard them altogether—a solution that, however, would simply go against the fact that such idealizations are not only tolerated but are successfully employed in the formulation of laws and principles that have high epistemic value in science. Nevertheless, I argue that in this case the puzzle can indeed be solved if we understand the proper role of such idealizations.[14] Namely, I submit that idealizations that cannot be true are not meant to be false or approximate representations of natural phenomena but rather play a

[9] See Morrison (2015, 85–6). Again, I take Morrison's 'abstraction' to be a form of idealization.

[10] By 'x is possibly true of nature', I mean that 'x is physically conceivable'. An assumption such as infinite population is incompatible with known laws of nature, whereas Maxwell's model of the ether is imaginary yet compatible with mechanical laws (see Maxwell 1861/62, 486). For the notion of physical conceivability, see Massimi 2019.

[11] Among various proposals, Weisberg (2007) distinguishes between three kinds of idealizations: Galileian idealization, minimalist idealization, and multiple-models idealization. I submit that there can be different, meaningful ways of classifying idealizations.

[12] For an attempt to reconcile idealizing with factive explanation, see Sullivan & Khalifa 2019. For an example of non-truth-based notion of scientific knowledge, see Potochnik (forthcoming).

[13] What about possibly true assumptions that seem to play a necessary explanatorily role such as the assumption that 'like begets like' in evolutionary game theory models? As possible representations of phenomena (objects or causal dependencies embodied in those objects), I take such assumptions to be open to correction and refinement. Assume that this is not the case because they map onto some aspect of the phenomenon or causal connection embodied in them. In this case, I suspect that we should not regard such assumptions as idealizations—they would be (approximately) true representations of objects or processes. For a different view, cf. Potochnik 2013.

[14] For the purposes of the paper, I will focus on idealizations (2) and set aside idealizations (1).

fundamental second-order role that allows us to come up with increasingly more unified and accurate explanations. As I will show in detail, this new role will allow us to understand why some idealizations are as such (not just temporarily or heuristically) necessary and compatible with a broad understanding of scientific realism. Moreover, the proposed approach improves not only on the heuristic and fictionalist views but also on recent attempts to reconcile idealizing with factive understanding.

The solution I will offer is inspired by Kant's doctrine of the regulative use of reason and its ideas. [15] For Kant, regulative ideas are not ~~just~~ false representations or misrepresentations. They are rather "necessary concepts of reason to which no congruent object can be given in the senses" (A327/B383). More specifically, they are concepts of totalities or wholes that overstep the possibility of human cognition but have an "indispensably necessary regulative use" (A644/B672). Examples include, among others, the idea of a fundamental power in psychology (A649/B691), the ideas of pure chemical elements ("pure water", "pure earth", and "pure air"; A646/B674), and the ideas of the totality of empirical causes, material and temporal parts, and contingent existences (A412–5/B439–43). Following Kant, I propose to regard contemporary idealizations (2) as idealizations in Kant's regulative sense of the term. I wish to clearly isolate such idealizations from the rest of ideal representations used in science and call them 'regulative idealizations' (RI).

Isolating this type of idealizations from the others is promising for two reasons. (i.) As we saw, this kind of idealization is not just a temporary expedient or calculation tool, it is rather key to formulating scientific laws and principles. As Kantian regulative ideas, they seem to play a necessary role in science that cannot be simply dispensed with (contra the heuristic view). (ii.) As clearly put by Morrison, what I call here idealizations (2) are indispensable because they seem to play some specific role in our cognition. Admittedly, such a role remains highly mysterious at this stage, and it is not clear whether it is compatible with realism—one may be tempted to embrace non-realism to account for their epistemic success. I will show that Kant's approach has internal resources to elucidate the epistemic validity of regulative idealizations in a way that makes them not only compatible with truth but, in some cases, a condition thereof (contra the fictionalist view). To do so, let's take a closer look at how idealizations are said to contribute to cognition in the literature.

4. The relation between idealization and truth

In this section, I will focus on the relation between idealization and truth. As we saw, fictionalist readers recognize the necessary epistemic value of idealization while, however, abandoning veritism. One prominent reading of the role of idealization in understanding has been proposed by Potochnik (Potochnik 2017). I take her to distinguish two fundamental dimensions of explanation: the ontic and the epistemic dimension. An explanation ('A because B') requires that we identify what explains something else—we need to identify the B that is causally responsible for A. To this aim, accuracy and truthfulness is all that matters. But this is only one side of explanation: the subject must also be able to understand and communicate 'A because B' (see, e.g., Potochnik 2017, 125–126). That's where, according to Potochnik, idealization becomes relevant to us. Since our real-world is complex and human understanding is constitutively limited, idealizations give a key contribution to our grasping of natural phenomena. As it were, their falsity allows the understanding to get a hold of complex

[15] Although Kant does not directly reference the very term 'idealization', I think it is plausible to take regulative ideas as the Kantian analogue of (a certain class of) contemporary idealizations.

phenomena by introducing order and patterns. Crucially, from these insights she concludes that understanding is not just about truth:

> Deviations from truth can positively contribute to understanding—that is, less truth can, in the proper circumstances, lead to greater understanding than would more truth … Thus, in science, truth play only a supporting role. (Potochnik 2017, 103)[16]

I subscribe to the idea that (i.) idealization promotes understanding in a positive way, but I resist the claim that (ii.) idealization moves us away from truth. I wish to argue that, in the case of regulative idealizations (RI), (ii.) follows from two tacit and avoidable presuppositions: namely, (a.) that idealization is, or is part of, an explanation; (b.) that idealization has epistemic value based on its descriptive power. The trade-off between the epistemic and ontic dimension of explanation results from presupposing these two premises. Following Kant, I propose a different account of the relation among idealization and truth. Idealization does promote 'understanding' (in both the contemporary epistemic meaning of the term and, as we will see, its Kantian technical meaning), but understanding should be conceived of as a means for truth. Let's first see how Kant describes such relation.

In Kant's system "understanding" (*Verstand*) is a technical term. It does not refer to any kind of epistemic achievement, but to the intellectual faculty that has the closest relation to objects. Kant defines it as the "faculty of concepts" (A160/B199)[17] and, as the faculty that conceptualizes objects, it is also responsible for the empirical investigation of nature and truth. Ideas, for Kant, do not belong to the "understanding" but to "reason" (*Vernunft*). By contrast with the understanding, reason is a second-order intellectual faculty: the faculty of principles that serve for "comprehension" (A311/B367). As we saw, ideas of reason go beyond the possibility of experience. As such, no object can correspond to them. That means, for Kant, that ideas cannot be related to objects: "if their significance is misunderstood and they are taken for concepts of real things, they can be transcendent in their application and for that very reason deceptive" (A643/B671). Their proper or immanent use, instead, consists in relating them to the operations of the understanding itself. As he puts it:

> Reason never relates directly to an object, but solely to the understanding and by means of it to reason's own empirical use, hence *it does not create any concepts (of objects) but only orders them and gives them that unity which they can have in their greatest possible extension*, i.e., in relation to the totality of series; the understanding does not look to this totality at all, but only to the connection through which series of conditions always come about according to concepts. Thus reason really has as object only the understanding and its purposive application. (A643–4/B671–2; my emphasis)

The second-order function of reason may incline us to think that ideas are only a desideratum of cognition: something to pursue only to strengthen or improve our body of cognitions. I take Kant to mean something more substantial. For Kant, ideas essentially guide the actions of the understanding—they bestow on it a unity that we could not otherwise achieve and by doing so, they make their use "purposive". Kant continues:

> They (ideas) have an excellent and *indispensably necessary regulative use*, namely that of directing the understanding to a certain goal respecting which the lines of direction of all its rules converge at one point, which, although it is only an idea (*focus imaginarius*)—i.e., a point from which the concepts of the understanding do not really proceed, since it lies

[16] One may object that Potochnik does not take an anti-realist stance in her book—as she puts it, her view "does not entail anti-realism" (119). However, she denies that the aim of science "is best articulated as truth" (ibid.). As a result, I take her view to call into questions notions of realism defining epistemic success in terms of (approximate) truth.

[17] References to the *Critique of Pure Reason* follow the standard A/B pagination, citing Kant (1998).

> entirely outside the bounds of possible experience—nonetheless still serves to *obtain for these concepts the greatest unity alongside the greatest extension*. (A644/B672; my emphasis)

Ideas are "indispensably necessary" to obtain greatest unity and extension of the concepts of the understanding. In other words, they maximally unify and extend our cognition. Kant calls such unity "systematic" (A645/B673). In itself, Kant says, "systematic unity" is only a projection, i.e. a vanishing point or *focus imaginarius* for our investigation. But crucially, systematic unity is also a "touchstone of truth" for the rules of the understanding, namely a necessary condition of empirical truth (A647/B675). As he puts it, "this unity helps to find a principle for the manifold and particular uses of the understanding, thereby guiding it even in those cases that are not given and making it coherently connected" (ibid.; see also A651/B679, A680/B707).

I take Kant's system as a springboard to draw up a new map of the relation between idealization and truth. Idealization is not simply a way to facilitate 'understanding' (in both the contemporary and Kantian sense). It is rather what guides our empirical investigation of nature, which is on the other hand concerned with truth. If that's correct, the ideas of reason can guide our empirical investigation and, *by doing so*, contribute to truth. There seems to be no real conflict between ideas and truth in Kant's account. While this new account seems promising for our purposes, its details still need to be fleshed out. In the next section, I will briefly elucidate how reason's ideas are supposed to guide our investigation in its relation to objects, namely truth.

5. Kant's regulative idealization and truth

How do idealizations promote truth according to Kant? This is not an easy question to answer. Ideas do have representational content: they represent chemical elements as pure, the series of causes as complete, psychological powers as fundamental, etc.[18] But if they cannot be related to objects, such representational content cannot be interpreted in standard descriptive terms. Better: the descriptive relation between an idea and its supposed object cannot ground the success of their use precisely because there may be no object the ideas describe. In other words, ideas have representational content, but their successful use, as it were, cannot rely on this content. If that's correct, their proper use must not be descriptive.[19]

As Kant puts it, ideas must be used *regulatively*—i.e. as rules for the systematization of empirical investigation. As Kant puts it:

> They are merely thought problematically, in order to *ground regulative principles of the systematic use of the understanding* in the field of experience in relation to them (as heuristic fictions).[20] If one departs from this, they are mere thought-entities, the possibility of which

[18] Kant's views about how ideas 'represent' (or 'present') are complex and cannot be summarized satisfactorily here. For my purposes, however, it should suffice to note that ideas can never be directly "projected in an image" or "given *in concreto*" (A328/B384–5). However, they can "represent other objects to us", namely empirical objects, systematically (A670/B698). We can, for example, investigate chemical elements by considering them *as if* they were derived from ideal pure elements. The characterization of the representational content of ideas is, however, purely analogical and is not meant to give a determination of a transcendent object (see, e.g., A674–5/B702–3). Cf. also Kraus on the regulative use of ideas as an "act of presentation" (2020).

[19] This is not to say that nothing can correspond to ideas, or that there cannot be other ways in which ideas refer to something. Thanks to X for pressing me to clarify this.

[20] By "heuristic fictions" I take Kant to mean that ideas are not representations of real objects, but, in a way, mental representations that play a role in our investigation of nature. One should not take this characterization to

> is not demonstrable, and which thus *cannot be used to ground the explanation of actual appearances through an hypothesis*. (A771/B799; my emphasis)

Such rules have a fundamental role in guiding the subject with respect to phenomena and their explanations. Crucially, that does not mean that ideas are themselves explanations. Why is this the case? I cannot offer here a general theory of explanation in Kant, but it should suffice to say that, for him, explanations have "grounds" (e.g. B115) or "physical grounds" (e.g. A690/B718) and must be connected with possible experience (A770/B798). Now, ideas are in themselves mere thought entities and since they do not describe possible objects, they cannot be used as plausible hypotheses in the explanation of certain explananda.[21] Instead, they should be strictly used as *rules* for the inquirer and, more specifically, as rules for the systematization of empirical investigation. What does this mean? Aren't grounds of explanation all we need to further science? That explanations must be physically grounded does not commit Kant to a fully-fledged ontic conception of explanation, i.e. the view that explanations are entities in the world.[22] For while explaining must be grounded in facts, it is not a purely ontic matter. Rather, an explanation is, first and foremost, a representation. We typically favour explanations that are general, unified, but also complete and precise—in a word, *systematic*.[23] I take Kant's regulative ideas to spell out the constrains of the explanations that we ought to seek in experience. For example, we systematize different causal relationships between materials according to our ideas of chemical elements, or we look for ever more fundamental empirical powers according to the idea of fundamental power.[24]

In other words, ideas function as "canons" (A329/B385) or "standards" (A675/B703) that do not tell the subject how an object is constituted but rather how she *ought* to seek after explanations with respect to certain phenomena. But what exactly do such rules demand? I suggest that, for Kant, rules of reason have both a (i.) *constructive* and a (ii.) *normative* dimension.[25] (i.) First, they create the unitary domain or conceptual space within which rational inquirers can look for empirical explanations of certain given phenomena. By following the idea of fundamental power, we are not determining any object—we are rather establishing a unifying proxy that allows to investigate "the constitution and connection of objects of experience in general" (A671/B699). (ii.) Second, they have the normative function of representing a canon against which we can assess the particular explanations we identify—they allow us "to assess and measure the degree and the defects of what is incomplete" (A329/B385). For example, by using ideas as standards of explanation, we look for and

imply that Kant subscribes to the 'heuristic view' presented above (which emphasizes the instrumentality and dispensability of idealizations).

[21] Even more explicitly, Kant also says that if ideas are used as hypotheses they would be "no explanation at all" (A772/B800).

[22] See Bokulich (2016) for a critique of this view.

[23] For Kant the systematic in cognition is "its interconnection based on one principle" (A645/B673). Weisberg (2007) lists values such as completeness, simplicity, causal relevance, accuracy, and generality as representational ideals pursued by idealizations. I take them to be plausible ways to articulate different aspects of Kant's systematicity in contemporary terms (although an in-depth discussion of such values is not possible here).

[24] A Kantian explanation can very well contain generalizations, fictional representations, or even idealizations (1). All these representations, however, are meant to contribute to our understanding of nature directly or exploratorily. That's not the case of regulative idealizations. If they contribute to truth, as Kant argues, they can only do it in indirect and indeterminate ways (A665/B693). Conversely, since we do not know whether ideas are possible (A771/B799), they cannot be used to map onto actual possibilities. Using them as such would fail any scientific standard.

[25] My view benefits from and expands two recent normative approaches to Kant's ideas of reason: Massimi's 'I-Rule' interpretation of the ideas of reason (Massimi 2017) and Kraus's normative reading of the idea of the soul (Kraus 2020). I develop my reading of Kant's regulative use of reason in general in XXX.

discover ever more fundamental powers in psychology (A649/B677), or we use the assumption of pure elements in chemistry "in order to explain the chemical effects on materials" (A646/B674). Since (i) and (ii) lead us to discover new empirically grounded explanations, or to systematize the ones we already found, it also follows that regulative idealization spurs the discovery of empirical truths. Finally, since ideas only *demand* systematic explanations without presupposing that they will be found, they do not induce us into errors. While we may not find systematic explanations, or they may not be available as such, we are still justified in looking for them.

To sum up, Kantian idealizations are representations that contain the instructions to guide epistemic agents in a complex, non-ideal world. For example, by telling them to look for ever more fundamental powers, or for a complete series of causes. By doing so they do not establish tentative descriptions of objects. Rather they enable and promote scientific explanations—"they initiate and continue" our actions of empirical investigation (A508/B536). More specifically, as I will further show below, they have the twofold function of (i) creating the domain of explanation and (ii.) of being standards for the assessment of explanations. As such, they are not antithetical to truth but play a crucial role in our very quest for it.

6. Example: Hardy-Weinberg equilibrium (HWE)

At this point, it is instructive to introduce an example that is often discussed in the idealization debate and that, I contend, can be interpreted along the Kantian lines here proposed: the Hardy-Weinberg equilibrium (HWE). Such principle (or model) occupies a central role in biology (its introduction initiated the field of population genetics as we know it) and yet it presupposes several unrealistic assumptions to be formulated, including infinite population. The principle is described by the two equations below (*Fig. 1*) and debunks the view that a dominant allele would tend to increase in frequency in subsequent generations. HWE states that the genetic variation in a population will remain constant in the absence of disturbing factors. More precisely, HWE is the state of the genotypic frequency in the simplest case of a single locus with two alleles after one generation of random mating without disturbing factors, such as mutation, selection or other forces, in an infinite population that experiences no genetic drift (i.e. change in frequency of an allele due to random sampling in a finite population).

$p^2 + 2pq + q^2 = 1$	p = frequency of the dominant allele in a population
$p + q = 1$	q = frequency of the recessive allele in a population

Fig. 1. HWE equations

The HWE is a paradigmatic case for the role of idealization in the life-sciences and many different approaches have been advanced to explain its relevance in biology. On the one hand, HWE is an elegant principle that has vast application and theoretical significance. At the same time, it is a highly idealized model. Of course, there cannot be infinite populations in nature and as a result the model does not describe any existing population. Moreover, the relative frequencies that HWE allows us to calculate lose all meaning when the population size is infinite. As pointed out by Abrams (2006), relative frequencies (other than zero and one) are only defined for finite populations—there is no determinate ratio of infinitely many As to infinitely many Bs.[26] But if this is correct, one may think that HWE is a mistaken principle and biologists should give up talk of infinite populations, or at best, use the HWE as a mere calculation device (see Edwards 1977; Abrams 2006). With few exceptions, this is not what biologists and philosophers of science think about HWE. They agree that despite (or, in some cases, in virtue of) its idealized nature, the principle does play a key role in our understanding of biological population (e.g., Morrison 2015; Strevens 2019). So why do they think that is the case? And how can the regulative approach help us understand its role?

I will first introduce two influential readings of the epistemic role of HWE and then present my regulative reading (in the following section). The first reading (by Strevens) belongs to the heuristic/factivist tradition; the second one (by Potochnik) to the fictionalist/non-factivist strand of literature. My reading differs from both traditions. It agrees with factivism that explanation cannot be directly provided by an in-principle-false assumption, but also endorses the non-factivist tenet that idealization (here and hereafter meant as idealization (2) or regulative idealization (RI), unless specified) plays a positive key role in our understanding of explanation.

Let's start with Strevens's reading. According to Strevens, the infinite population idealization is obtained through the following asymptotic process (*Fig. 2*):

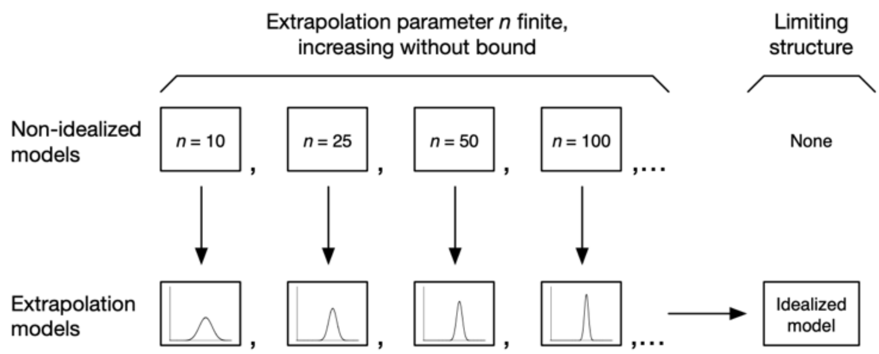

Fig. 2. Asymptotic extrapolation according to Strevens (2019). Reproduced with permission from Springer Nature.

Fig. 2 visualizes the process of asymptotic extrapolation on a stochastic population genetics model with population size n (extrapolation parameter). First, we assign a fixed value to n (in this case, 10) and thereby we obtain a new model in which n cannot change (upper row). Then we extrapolate a model that has no reference to n by replacing absolute numbers with relative frequencies (lower row). The variance of the distribution represents genetic drift, i.e. the chance due to random sampling that allele frequencies change from one generation to the next. Extrapolation can be done for any value of n. The idealized model is created by using asymptotic reasoning: as n increases indefinitely, models converge towards a "limiting structure" (Strevens 2019, 1725). In this case, the limiting structure is a distribution that assigns probability 1 to the frequency of the next generation. Importantly, the idealized model is not simply obtained by assigning an infinite value to *n*. Rather, it is obtained via a sequence of extrapolation models that only represents relative frequencies. This process allows Strevens to avoid reliance on the existence of a model with an infinitely large population. In fact, there is

[26] See Strevens (2019, 1721).

no structure in the model that directly represents infinitude. What is then the function of asymptotic idealizations such as infinite population according to Strevens? Answer: idealizations play an important negative role: “they underline a factor’s explanatory irrelevance” (Strevens 2019, 1727). In other words, idealizations are shorthand for indicating that a certain factor should be considered irrelevant to the phenomenon that is to be explained. In this case, the idealization merely stands for the fact that a certain disturbing factor dependent on the size of the population (genetic drift) is “irrelevant to the outcome of the evolutionary explanandum” (1728).

While Strevens’s solution elegantly avoids reliance on existing infinite populations, I agree with Potochnik (2017) that it falls short of explaining the role HWE plays for our understanding. If the only function of HWE is to declare genetic drift irrelevant, HWE would seem nothing but a calculation tool. But why should we extract such an ideal model when genetic drift is indeed relevant to any finite population? The creation of an extrapolation space seems rather *ad hoc*. On the contrary, Potochnik thinks that idealizations “indicate the nature of a factor’s relevance to the focal causal pattern” (Potochnik 2017, 52). Potochnik’s explanation of how this works relies on a general account of representation. More specifically, idealizations are taken to bear key similarities to the target system they shed light onto. On her account, ideal (mis)representations are supposed to share “functional similarities” with some features of represented phenomena (ibid., 53), where functional similarities are defined as “similarities in causal role or behavior” (ibid.).

While promising with respect to idealization in the broad sense, I am less confident this approach is successful in vindicating the role of what I have called in this paper regulative idealizations (RI). As I see it, the problem is that such functional similarities between an idealized model and the target system can be established only if the idealization (or some feature of it) is a representation that *can* correspond to some existing object or feature of an object. For instance, effective population size N_e (the example Potochnik uses) is an idealization that shares with the actual population in question the same magnitude of random genetic drift. That similarity between N_e and actual population with respect to genetic drift can be used to shed light on some focal causal pattern. The case of the infinite population idealization, however, is different. In this case, there is no possible (i.e. physically conceivable) existing population that shares the same features (for instance, the same null genetic drift) of an infinite population. To say that there is a similarity between infinite and actual populations in terms of relative frequencies is also not particularly promising since relative frequencies have no meaning for infinite populations. It is thereby unclear what the referent of the similarity really is in this and similar cases. If it is a real referent (a real entity, property, or process, as Potochnik seems to imply), this reading results in the reification of an idealization that, by definition, cannot be actualized. If there is no real referent corresponding to the idealization, then it is unexplained how idealizations are meant to *positively* represent causal patterns that are relevant to the explanation of given phenomena.

7. Regulative idealization

As briefly anticipated, I agree with both key intuitions behind the two approaches I have illustrated in the previous section. I agree with Strevens that idealizations cannot have a direct representational role in explanation, but I also think, with Potochnik, that there is much more to them than mere elimination of irrelevant factors in the explanation—what is then the positive role of idealizations?

My claim is that idealizations play a positive role in science inasmuch as they have—to use Kant's terminology—an essential regulative function. Idealizations are not to be conceived of as mere fictions or as hypotheses. Rather, they are necessary rules governing the (i.) construction and (ii.) assessment of scientific explanations. What does this mean in the present context? Take the example of infinite population idealization. I think Strevens is right in pointing out that we must not presuppose an actual infinite population to construct this model—the idealized model is rather derived through an asymptotic derivation. In Kantian terms, this assumption is a "*focus imaginarius*" or imaginary standpoint towards which our empirical findings are supposed to converge. But I think Strevens undermines the function of such ideal standpoint by saying that it just cancels out a factor from being explanatory relevant. The assumption (as a mere rule) is precisely *what* creates the space where the asymptotic approximation can take place. It is only by projecting an idea of infinite population that empirical, finite population can be said to approximate to a unitary value of frequency at the increase of *n*. More precisely, this happens through the following two steps:

(i.) **Construction**. The introduction of the idealization has a first important result: it creates what Kant calls unity of cognitions. Instead of an unruly heap of empirical results, biologists are provided with a template that allows them to analyze gene and genotype frequencies in all finite populations. That is clearly a *positive* role. The introduction of the infinite population idealization grounds an ideal model (HWE) that acts as the founding block and starting point of any population genetics analysis. In other words, the introduction of the idealization creates the domain within which biologists can construct explanations of any finite population frequencies. On the other hand, we should be wary of such ideal unity. To use, again, Kantian terminology, the unity grounded by idealizations can only be regulative. As clearly recognized by biologists, HWE is not a law in any standard (or, to stick to Kantian terms, 'constitutive') sense of the term. Even when the expected values are confirmed, this should be taken as non-rejection of HWE rather than a proof that a given locus really exhibits HWE (see, e.g., Waples 2015). HWE is instead a 'rule' giving us a standard or canon that we can use to assess particular explanations.

(ii.) **Assessment**. HWE does not give us fixed results—it is rather a reference point that biologists can use to compare empirical data. In the simplest cases, HWE gives approximations that need to be tested using the actual genotypic frequencies. Such conditions are realized when all assumptions (including binomial sampling of gametes and panmixia) are roughly met. The standard method consists in generating expected genotypic frequencies according to HWE and then compare these with observed genotypic frequencies using a chi-square or related test (Waples 2015). If no significant deviation is detected, the actual data are said to approximately follow HWE. But crucially, this is not the main function of HWE. HWE gives rise to the most important findings when it is used as a standard "against which the effects of selection, linkage, mutation, inbreeding and chance can be detected and estimated" (Mayo 2008, 249). In other words, when deviations from HWE are significant and such deviations signal, through a deficiency of either or both homozygotes, or of heterozygotes, etc., the type of explanation that should be sought (ibid., 254). By using HWE as a standard for comparison, it is therefore possible to detect and measure the forces that disrupt the equilibrium. Biologists then elaborate more unified and accurate explanations of the results, thereby deepening our understanding of genotypic frequencies in all population-genetical investigation.[27]

[27] For an instructive illustration of the assessment procedure in HWE, see Table I in Waples (2015). The table relates the observed deviations from the equilibrium to a series of possible causes and explanations. For example, factors that commonly produce deficiencies of heterozygotes are null alleles, self-fertilization, the Wahlund effect and positive assortative mating (see ibid., 11).

The proposed view not only reconciles key insights coming from the two dominant approaches to the problem of idealization—it also advances recent, alternative attempts to explain the contribution of idealizations to factive understanding. Rice (2016), for example, argues that idealizing produces factive understanding in a variety of contextual ways, most importantly by providing modal information about counterfactual relevance. Pincock (2021) interprets idealizations as having a non-representational function that consists in signaling the commitment to the whole derivation of the explanation generated by the idealization via truths underlying it (i.e., the idealization is partially true). Finally, Lawler (2021) proposes an "extraction" view according to which idealized models have epistemic value since they enable the extraction of propositions that contribute to our understanding. These approaches are promising and contain insightful analyses. In my opinion, however, they do not offer fully satisfactory answers to two concerns. First, since they do not distinguish between idealizations that can or cannot be true, they either attribute descriptive power to idealizations that cannot be true (similarly to fictionalists; as in Rice 2016 and Pincock 2021),[28] or deny that any idealization can figure into the content of the understanding (something that is arguably justified with respect to idealizations that can be true, such as approximations or simplified representations; as in Lawler 2021). Second, the Kantian approach here proposed offers a determinate answer to why and how some idealized models are epistemically valuable. Regulative idealization contributes to the way we obtain factive understanding. And it does so in a specific way, by telling scientists how to construct and improve explanations in different contexts. For this reason, regulative idealizations are to be regarded as rules that are indispensable to the extraction or derivation of factive explanations.[29]

If this is correct, regulative idealization is compatible with and beneficial to realism, as I characterized it. For if truth is a necessary component of realism (albeit not necessarily a sufficient one), and regulative idealizing provides both the conceptual space and the normative standards to deepen factive explanation, it follows that regulative idealization also promotes the acquisition of truthful propositions based on facts. As suggested by Rice (2016), it seems plausible to assess the contribution of regulative idealization to truth as a global matter, so that local employment of falsehoods can still have a positive effect on the overall truth-balance, so to speak. My account adds some further qualifications to this line of thought. Since regulative idealizations are not themselves parts of explanation, their being false does not count as 'less truth' even at the local level of explanation. Further, since these idealizations play a constructive and normative role, they do not just promote the mere maximization of truths, but rather the acquisition of epistemically justified and explanatorily relevant truthful propositions. As a result, my account is amenable to views that qualify truth as part of the aim of science or reconfigure such an aim in epistemic terms (see Bird 2007).

The compatibility between regulative idealization and realism comes with some caveats. I suggested that regulative idealizations are idealizations that cannot be replaced, and in virtue of not being replaceable they offer an indirect, yet indispensable contribution to truth. A first caveat concerns how to practically identify indispensable regulative idealizations. A useful criterion is to isolate those models that cannot be made more accurate by adding correcting factors (as suggested by Morrison 2015). But this would still be an insufficient criterion for we cannot exclude that a breakthrough in a certain field makes more realistic explanatory models available. Second, if we do not have a definitive criterion of indispensability, some idealized

[28] By either presupposing that idealized models contain true modal information about the world (Rice 2016) or that they are partially true (Pincock 2021).

[29] I am thankful to an anonymous reviewer for pushing me to clarify these points.

models that we currently deem necessary may turn out to be incompatible with realism. A fully-fledged account of the compatibility between regulative idealization and realism should answer complex questions about the evolution and progress of science that I cannot address here. Let me briefly suggest, however, that, the lack of arguments for the absolute indispensability of regulative idealization is to be expected. Scientific investigation is a historical and situated endeavor. As far as we know, some idealized models are necessary to get epistemic access to explanatorily valuable information. We have no reason to think that we will dispense with them in the future, although we cannot exclude this possibility. Importantly, such a relativized notion of indispensability still allows us to maintain a distinction between rules that, albeit replaceable in principle, currently play an indispensable epistemic function (regulative idealizations) and idealizations that ought to be de-idealized or corrected.

Finally, let me briefly note that although the present proposal is inspired from Kant's philosophy and his idiosyncratic account of science, similar views have recently found support in psychological research. Kon and Lombrozo have shown that ideals, although rarely (or not to be) found in real scientific practice and everyday thinking, still play an important role in explanation. Their research "suggests that even if a simple, exceptionless pattern describes some explanatory ideal that is rarely realized, the pursuit of this ideal could spur meaningful discoveries" (2018, 1940). In their experiment (*Fig. 3*), they asked participants to classify items into one of two categories. They were prompted to 'explain' (by seeking an ideal rule) or simply write down the category membership of twelve exemplars. Only two rules could be used to categorize items in ideal and non-ideal world conditions: an easy to find 'worse rule' with a limited success rate in capturing category membership, and a 'better rule' that was more difficult to find but also more successful in categorizing items. Results suggested that when prompted to identify ideal patterns, subjects were more likely to identify the better rule of assortment than when operating without an ideal assumption (in both ideal and, importantly for this paper, non-ideal world conditions). The reason the authors identify for this phenomenon is that idealizing supports discovery: idealized models give us a starting point and a standard to find and assess unregular and chaotic patterns in a non-ideal world. I suggest that RI plays precisely this role in empirical investigation and that Kant's account of science provides a philosophically rich framework for investigating idealization as a necessary condition of empirical discovery and explanation.

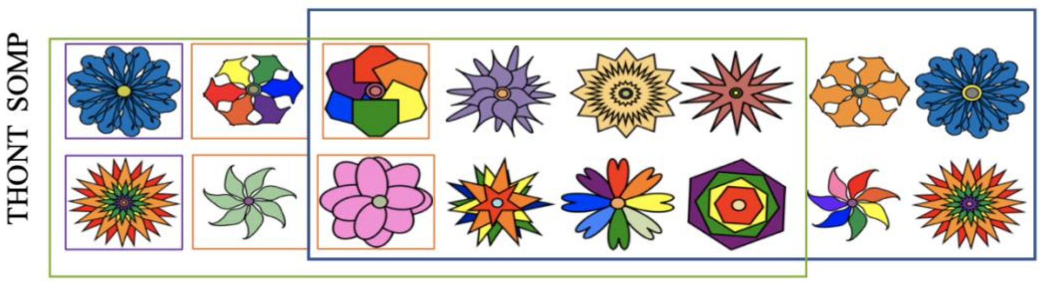

Fig. 3. Kon and Lombrozo's experiment. The better rule is that SOMP flowers have two concentric circles in their centres while THONT flowers have one circle in their centre. The worse rule is that the petals of SOMP flowers are mostly one colour, whereas the petals of THONT flowers are mostly rainbow coloured. The blue box contains the items in the ideal scenario; the green box contains the items in the non-ideal scenario (see ibid., 1941). Reproduced by permission from *Cognitive Science*.

8. Conclusion

I wish to conclude by clarifying the role of RI in the quest for scientific truth. If the analysis so far is correct, these idealizations do not stand for or describe real objects but are rather rules for the empirical investigation of nature. This characterization has important implications. As a ‘rule’, an idealization does not tell us what an object is—strictly speaking there is no HWE equilibrium as there is, in Kant’s philosophy of nature, no existing “pure water” nor “fundamental power”—but rather tell us how we ought to investigate empirical objects.

Although the rule itself is not an explanation, it imposes on cognizers epistemic constraints on explanations. This happens through two steps. First, it gives a set of phenomena a common referent point or starting point for explanation. That is the constructive function of idealization. Employing HWE is the starting point for all population-genetical investigation. In Kantian terms, it gives ‘systematic unity’ to our particular cognitions as a “*focus imaginarius*”. Second, idealizations are not explanations of phenomena, but rather normative standards for the assessment of explanations. It would be misleading to say that HWE *as such* ‘explains’ anything. It furnishes instead a standard to which we can compare collected data. By doing so we further the unity and accuracy of our explanations. More specifically, we identify cases where deviations are small and thereby do not reject HWE, and cases that deeply challenge HWE and thus lead us to new findings.

I submit that Kant is right in pointing out that idealizing is a condition or a “touchstone” of truth —not because it gives ideal truths about a non-existent world, but rather because it gives us rules on how to promote investigation and look for more unified and accurate explanations. As Kant puts it:

> One cannot properly say that this idea is the concept of an object, but only that of the thoroughgoing unity of these concepts, insofar as *the idea serves the understanding as a rule*. Such concepts of reason are not created by nature, rather *we question nature according to these ideas, and we take our cognition to be defective as long as it is not adequate to them*. Admittedly, it is hard to find pure earth, pure water, pure air, etc. Nevertheless, *concepts of them are required* (though as far as their complete purity is concerned, have their origin only in reason) *in order appropriately to determine the share that each of these natural causes has in appearance*. (A645/B673–A646/B674; my emphasis)

The function of regulative idealizations is to guide the actions of the understanding. Their concepts are *required* to determine the causal role of empirical phenomena. By following them as rules of inquiry, explanations can be constructed and improved systematically to approximate truth. If this is correct, the regulative conception of idealization here suggested solves the puzzle of idealization we started from: (i.) at least some idealizations, regulative idealizations (RI), are necessary, ubiquitous assumptions in science; and (ii.) their role is not only compatible, but beneficial to the quest for truth.

Acknowledgements

I am grateful to the North American Kant Society for awarding an earlier draft of this paper their Wilfrid Sellars Prize. Special thanks to Angela Potochnik and Eric Watkins for providing me with insightful and thought-provoking comments on the draft within the context of the VNAKS “Choose your Ideal Commentators” Series in October 2022. Earlier versions of the paper have been presented at the workshop “Metaphysics as Modeling. Contemporary and Kantian Issues” (Trier, October 2021), the British Society for the Philosophy of Science Annual Conference (Exeter, July 2022), and the International Conference of the German Society for Philosophy of Science (Berlin, August 2022). The paper has greatly benefited from discussion with the audiences, especially with Mazviita Chirimuuta, Brigitte Falkenburg, David Hommen, James O'Shea, Markus Schrenk, and Adam Toon. I am very thankful to Kristina Engelhard and Michela Massimi for their constructive feedback on the ideas for this paper. Three anonymous reviewers for Studies in History and Philosophy of Science provided valuable and helpful comments. All errors are my own.

References

Bird, A. 2007. "What Is Scientific Progress?," *Nôus*, 41(1), 64–89.

Bokulich, A. 2009. "Explanatory Fictions." *Fictions in science: Philosophical essays on modeling and idealization*, 91–109.

Bokulich, A. 2015. "Maxwell, Helmholtz, and the Unreasonable Effectiveness of the Method of Physical Analogy", *Studies in History and Philosophy of Science*, 50, 28–37.

Bokulich, A. 2016. "Fiction As a Vehicle for Truth: Moving Beyond the Ontic Conception", *The Monist*, 99 (3), 260–279.

Elgin, C. Z. 2017. *True Enough*, Cambridge, MA: MIT Press.

Fine, A. 1993. "Fictionalism", *Midwest Studies in Philosophy*, 18, 1–18.

Frigg, R. 2010. "Models and fiction", *Synthese*, 172 (2), 251–268.

Kant, I. 1998. *Critique of Pure Reason*, translated and edited by P. Guyer and A. Wood. Cambridge: Cambridge University Press.

Kon, E., and Lombrozo, T. 2018. "Seeking ideal explanations in a non-ideal world", In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Kraus, K. T. 2020. *Kant on Self-Knowledge and Self-Formation*, Cambridge: Cambridge University Press.

Lawler, I. 2021. "Scientific Understanding and Felicitous Legitimate Falsehoods," *Synthese*, 198, 6859–6887.

Massimi, M. 2017. "What Is This Thing Called "Scientific Knowledge"? Kant on Imaginary Standpoints and the Regulative Role of Reason", *Kant Yearbook*, 9, 63–84.

Massimi, M. 2018. "Perspectival Modeling", *Philosophy of Science*, 85 (3), 335–359.

Massimi, M. 2019. "Two Kinds of Exploratory Models", Philosophy of Science, 86 (5), 869–881.

Maxwell, J. C. 1861/2. "On Physical Lines of Force", reprinted in (1890) *The Scientific Papers of James Clerk Maxwell*, Dover: New York, 451–513.

Mayo, O. 2008. "A Century of Hardy–Weinberg Equilibrium", *Twin Research and Human Genetics*, 11 (3), 249–256.

Morrison, M. 2015. *Reconstructing Reality: Models, Mathematics, and Simulations*, Oxford: Oxford Studies in Philosophy of Science.

Niiniluoto, I 2017. “Optimistic Realism about Scientific Progress,” *Synthese*, 194 (9), 3291–3309.

Pincock, C. 2021. “A Defense of Truth as a Necessary Condition on Scientific Explanation”, *Erkenntnis*, https://doi.org/10.1007/s10670-020-00371-9.

Popper, K. R. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*, London: Hutchinson.

Potochnick, A. 2013, “Biological Explanation,” in K. Kampourakis (ed.), *The Philosophy of Biology: A Companion for Educators*, 49–65, Dordrecht: Springer.

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

Potochnik, A. (forthcoming). “Truth and Reality: How to be a Scientific Realist without Believing Scientific Theories Should Be True”, in K. Khalifa, I. Lawler and E. Shech (eds.), *Scientific Understanding and Representation: Modeling in the Physical Sciences*.

Psillos, S. 1999. *Scientific Realism: How Science Tracks Truth*, London: Routledge.

Rice, C. 2016. “Factive Scientific Understanding without Accurate Representation”, *Biology & Philosophy*, 31(1), 81–102.

Shech, E. 2018. “Infinite Idealizations in Physics”, *Philosophy Compass*, 13 (9), e12514.

Strevens, M. 2016. “How idealizations provide understanding.” In S. R. Grimm, C. Baumberger, and S. Ammon (eds.), *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*, New York: Routledge.

Strevens, M. 2019. “The Structure of Asymptotic Idealization”, *Synthese*, 196, 1713–1731.

Suárez, M. (ed.) 2009. *Fictions in Science: Philosophical Essays on Modeling and Idealization*, London: Routledge.

Sullivan, E. & Khalifa, K. 2019. “Idealizations and Understanding: Much Ado About Nothing?,” *Australasian Journal of Philosophy*, 97 (4), 673–689.

Vaihinger, H. 1911. *Die Philosophie des Als Ob*, Berlin: Verlag von Reuther & Reichard.

Waples, R. S. 2015. “Testing for Hardy-Weinberg Proportions: Have We Lost the Plot?”, *Journal of Heredity*, 106 (1), 1–19.

Weisberg, M. 2007. “Three Kinds of Idealization”, *The Journal of Philosophy*, 104 (12), 639–659.

Wimsatt, W. C. 1987. “False Models as Means for Truer Theories.” In M. H. Nitecki and A. Hoffman (eds.), *Neutral Models in Biology*, 23–55, New York: Oxford University Press.