Models, Idealisations, and Realism*

Juha Saatsi

27th October 2015

Abstract

I explore a challenge that idealisations pose to scientific realism and argue that the realist can best accommodate idealisations by capitalising on certain modal features of idealised models that are underwritten by laws of nature.

1 Introduction

This paper explores a challenge that idealisations pose to scientific realism. I will review the challenge before briefly assessing some recent analyses of idealised models that function as a foil and motivation for my response to the challenge. I will argue that the realist can best accommodate idealisations by capitalising on certain modal features of idealised models that are underwritten by laws of nature.

The idea that idealisations in some sense represent *possibilia* is common place. Typical idealisations—such as frictionless planes, point masses, isolated systems, and omniscient agents—are naturally thought of in terms of possible systems that are suitably related to some actual systems of interest. David Lewis, for example, thought that we can best make sense of the pervasive utility of idealisations in science in terms of possible worlds that are more and less similar to the actual world.

> [We find it much easier to tell the truth if we sometimes drag in the truthlike fiction, and when] we do, we traffic in possible worlds. Idealisations are unactualised things to which it is useful to compare actual things. An idealised theory is a theory known to be false at our world, but true at worlds thought to be close to ours. The frictionless planes, the ideal gases, the ideally rational belief systems—one and all, these are things that exist as parts of other worlds than our own. The scientific utility of talking of

*Acknowledgements: A version of this paper was presented at a workshop on Models and Inferences in Science in Rome. Thanks to the workshop audience, as well as James Fraser, Steven French, and especially Arnon Levy, for helpful comments.

> idealisations is among the theoretical benefits to be found in the paradise of possibilia. (1986, pp. 26–27)

Recognising that idealisations are naturally thought of in terms of possibilia is but a start, however. The spadework lies in properly accounting for the utility of such modal constructs.[1] What, then, is required to account for the utility of idealisations in science? Various questions present themselves here. I will focus on the following challenge to scientific realism, in particular: given that idealisations incorporate false assumptions about the way the world is, why are idealisations so important for making successful predictions and for coming up with powerful explanations?

I will examine this challenge in relation to idealised *models* in particular. (Often theoretical *laws* are also characterised as idealised. My focus is on idealised modelling assumptions other than laws.) This challenge differs from the standard arguments against scientific realism, deriving from a 'pessimistic induction' over past false theories, or the idea that theories can be underdetermined by evidence. These stock anti-realist arguments are typically framed in terms of scientific *theories*. It is interesting that the realism debate has been largely framed in terms of theories, even though in contemporary philosophy of science much of the focus has shifted from theories to models as the most fitting 'unit' of philosophical analysis. In as far as realism is primarily motivated by the impressive empirical success of science (culminating in novel predictions), it is typically models that provide or facilitate such success. Furthermore, according to a popular 'modelling view' of science, theories in a sense are nothing but families of models (unified by laws).

When we shift the focus from theories to models, anti-realists can find further ammunition from various kinds of inconsistencies that modelling practices exhibit. Often different models of one and the same phenomenon are mutually inconsistent with one another. Some models are even internally inconsistent. And many models incorporate assumptions that are at odds—sometimes radically so—with modellers' background beliefs. Such inconsistencies can be used to challenge the realist in as far as they indicate that various kinds of falsehoods are playing a bigger role in the production of the empirical successes that realists are inclined to think. If falsehoods can play a significant role in bringing about empirical successes, perhaps the role played by (approximate) truths is less significant than realists would have it? Perhaps the joint contribution to empirical success from idealising falsehoods, and whatever degree of (approximate) truth there is to a model otherwise, can be so entangled that we can make no sense of the realist credo that a model's empirical success is *due to* its 'latching onto reality'?

[1] A lot has been written about modal aspects of idealisations. I will not attempt to relate my point of view here to the broader context of the Poznań school and the verisimilitude literature, for example. See Niiniluoto (2007) for a review.

What follows is concerned with this kind of challenge, arising out of the indispensability of idealisations for modelling. I will start by fleshing out the challenge (§2), before briefly reviewing some philosophical analyses of idealisations (§3), paving the way for my own response to the challenge (§4).

2 Realism and idealisations: a challenge

Let us first try to get a good handle on the 'idealisation-challenge' previewed above. How exactly do idealised models challenge a realist attitude to science? Sorensen (2012) crisply (if somewhat provocatively) explains:

> Scientists wittingly employ false assumptions to explain and predict. Falsification is counter-productive in the pursuit of truth. So scientific realism appears to imply that idealisation would be worse than ineffective.
>
> The instrumentalist says the scientist merely aims at the prediction and control of phenomena. [...] Given that scientists are indifferent to the truth and often believe idealisations will promote prediction and control, the instrumentalist predicts that the scientists will idealise.
>
> Consequently, idealisation looks like a crucial experiment for philosophy of science. [...] Since scientists idealise, the instrumentalism prevails. (p. 30)

In other words, if realism is committed to the notion that science aims at truth, while anti-realists regard lesser aims of empirical and instrumental adequacy to be enough, then idealisations seem to speak against realism.

One might worry that this challenge to realism quickly evaporates in the light of obvious realist responses applicable to many (or perhaps even most) idealisations.[2] Consider various 'Galilean' idealisations, for example, that McMullin (1985) views as providing an argument *for* scientific realism, not against it. Take an idealised model of a gravitational pendulum, for instance. It incorporates various simplifying assumptions, such as the complete absence of air resistance, friction, and so on. But it does so in a way that readily suggests ways in which the model can be *de-idealised*, for example by simply adding further terms to the model's force function. McMullin rightly points out that a realist reading of idealised models best predicts and explains models' capacity to be thus de-idealised; therefore such idealisations arguably support (a suitably qualified form of) realism about these kinds of idealised models.[3]

[2]To be clear, Sorensen himself notes that idealisations only 'appear' to challenge scientific realism, and he does not endorse the instrumentalist conclusion in the offing. I will review Sorensen's reasoning in §3.

[3]See McMullin (1985):

> If the original model merely 'saved the appearances' without in any way approximating to the structure of the object whose behavior is under scru-

In general, a realist perspective on scientific modelling clearly has the wherewithal to account for the way in which various kinds of simplifications are *pragmatically* indispensable in the scientific study of systems of otherwise unmanageable complexity. Realist reading of theories suggests different ways of brushing aside complications that, according to our theory, make next to no contribution to the end result. After all, science is obviously *not only* in 'the pursuit of truth', even according to the realist; it is also in the pursuit of achieving actual results, mathematical tractability, predictions, effective control and manipulability, and so on. The different aspects of Galilean idealisations are ways in which the realist can anticipate deliberate 'falsifications' (typically simplifications) to contribute to the latter pursuits.[4]

This realist response does not answer the challenge completely, however, since some idealisations do not fit the Galilean mould. Philosophers of science have identified other, more radical idealisations in science, and the question remains whether some of these non-Galilean idealisations rather support instrumentalism about certain kinds of models. What should a realist say about 'uncontrollable' idealisations where no de-idealisation is in the offing? (See e.g. Batterman, 2005) What about idealisations involved in the so-called minimal models, such as the Ising-model? How should the realist accommodate these kinds of idealisations that seem altogether indispensable, going beyond the kind of broadly pragmatic convenience associated with Galilean idealisations? How can the realist account for the indispensable utility of such falsifications in modelling? This is one challenge that remains for the realist.

Moreover, even with respect to Galilean idealisations, there is further work to be done in clarifying the letter of the realist response. For example, is there a conceptual framework within which the utility of different types of idealisations can be accounted for in unified terms? Intuitively speaking, the realist response to the challenge from idealised models is to say that there is a sense in which an idealised model 'latches onto' reality in a way that is responsible for the model's empirical success. One challenge is to articulate this notion of 'latching onto reality' so as to capture the relevant features of models in a way that meshes with the realist intuitions. Call this the articulation-of-realism challenge. What does an idealised model 'get right' about its target system, such that it is empirically successful *by virtue of* getting those things right (and despite getting some other things wrong)?

> tiny, there would be no reason why this reversal of a simplifying assumption, motivated by the belief that the object does possess something like the structure attributed to it, would work as it does. Taking the model seriously as an approximately true account is what leads us to expect the correction to produce a verifiable prediction. The fact that formal idealisation rather consistently does work in this way is a strong argument for a moderate version of scientific realism. (p. 262)

[4] McMullin (1985) distinguishes three different types of Galilean idealisations.

This challenge of articulating how idealised models latch onto reality has been recognised in the vast literature on idealisations, and realists typically maintain that there is some principled sense in which predictive (as well as explanatory) success is due to models latching onto reality. This then underwrites the realist's epistemic commitment for regarding predictive success as a (fallible) indicator of models latching onto reality in this sense.

Philosophers have appealed to different conceptual and formal resources in spelling out this idea, ranging from accounts of verisimilitude, to partial structures and quasi-truth, to philosophy of language/logic, to philosophy of fiction.[5] I cannot do full justice to this rich literature here, but I will next briefly review a couple of recent analyses of idealisation as a foil for my own perspective. To prefigure: in my view these analyses fall short of properly accounting for the empirical success of idealised models. Providing a sense in which a model can get things right, while also getting things wrong, does not in and of itself account for how the falsehoods are *immaterial* for the empirical successes at stake, and how the empirical successes are *due to* 'getting things right'. After discussing these analyses of idealisations I will take steps towards a different (possibly complementary) account of idealisations that better serves the realist's need to explain the empirical success of idealised models. This requires reflecting more closely on what it takes to account for predictive success of a model that incorporates false assumptions. I will argue that such an account can turn on showing how a model's predictive success is *robust* with respect to variation in the false assumptions involved in idealisations, in the sense that these assumptions could have been different without undoing the predictive success. I will argue that it is this modal character of idealisations that can account for their utility from a realist perspective.

3 Some analyses of idealised models

Idealisations as suppositional. Recall Sorensen's presentation of the idealisation-challenge above. His own response to it is iconoclastic. Typically philosophers characterise idealisation as being essentially a matter of some sort of intentional introduction of distortion into a scientific model or theory, with different philosophers holding different views regarding the nature of such 'intentional distortions'. For example, such intentional distortion has been taken to be a matter of *indirect* assertion of something true (Strevens); *relativized* assertion of something true (Giere); *temporary* assertion of falsehood (McMullin); or assertion in the mood of *pretence* (Toon, Frigg, and various others). In contrast to these different ways of regarding idealisation as some sort of attenuated assertion, Sorensen views idealisations as *suppositional*, in analogy to suppositional premises in a con-

[5] See e.g. Niiniluoto (2007), da Costa and French (2003), Sorensen (2012), Toon (2012).

ditional proof.[6] That is, Sorensen's perspective on idealisation—drawing on philosophy of language and logic—regards it as a matter of 'simplifying supposition,' naturally free of any realist commitment. (Compare: a mathematician's supposition that 'there is a largest prime' for the purpose of *reductio ad absurdum* entails no commitment to finitude of primes.) In sum:

> Idealisation is constituted by supposition. Only simplified suppositions count as idealisations. The filters are psychological and methodological. Idealisers seek tractability, memorability, and transmissibility (rather like myth makers). (Sorensen 2012, p. 37)

Sorensen contends that we can thus assimilate idealisations with something that is already well understood by logicians—a supposition that initiates a conditional proof or reductio ad absurdum. Allegedly we thus have an 'off-the-shelf' model for analysing idealisations as a matter of propositions that are governed by well-understood rules of rational use; not ontologically committing, for well-understood reasons; and not in need of elimination. And all this arguably explains, at least in part, why scientists are so happy to idealise, and are not overly preoccupied with de-idealisation or the 'distance' of verisimilitude between idealisation and the exact truth.

While Sorensen's perspective may throw light on some idealisations in science, for various reasons I do not find the analogy convincing or illuminating in general. Even the simplest paradigmatic exemplars of Galilean idealisation, such as the ideal pendulum or a frictionless plane, seem to be fit-for-purpose for obvious reasons that have little to do with *reductio ad absurdum*, or purely conditional arguments. Analysing idealisations in terms of the status of propositions involved—suppositional vs. assertoric—also seems much too dichotomous and coarse-grained to capture relevant differences in the various kinds of idealisations and how they contribute to predictive success. (See e.g. McMullin (1985) for useful distinctions amongst different flavours of 'Galilean' idealisations, and Batterman (2005) for the distinction between these and 'non-Galilean' idealisations.) Furthermore, with respect to the articulation-of-realism challenge most importantly, it is wholly unclear why a realist account of a model's predictive success should in any way depend on whether or not the 'falsehoods' involved are intentional, as in the case of idealisations, or simply mistaken assertions about the target. In both cases we can consider the relationship between the target-as-represented-by-the-model, and the target-as-it-actually-is, in trying to account for the model's empirical success in terms of how it latches

[6] Schematically:

(P1) Suppose P.
(P2) From P derive Q.

(C) Conclude that if P then Q.

onto reality.[7]

Idealisations in the semantic view. According to the semantic view of theories idealisations are more of a piece with other approximations. The semantic view is touted as providing a unified account of science where models occupy a centre stage. Consider da Costa and French (2003), for example, who offer models in the sense of (quasi-formal) model-theory as an appropriate backbone to a 'unitary approach to models and scientific reasoning.' In particular, their model-theoretic meta-scientific framework is motivated as offering the wherewithal to capture idealisations and approximations in science by providing 'a more sophisticated concept of 'model' [...] which accommodates the essential incompleteness and partial nature of scientific representations.' (p. 5) In their 'partial structures' formalisation of the semantic view, idealisations (as well as other approximations) can be 'accommodated through the introduction of 'partial isomorphism' as the fundamental relationship—horizontally and vertically—between theoretical and data models.' (p. 102) Furthermore, the model-theoretic framework furnishes a notion of 'quasi-truth' that 'can be used to formally underpin the claim that idealisations are regarded *as if* they were true.' (p. 163) (See da Costa and French (2003) for details.) Moreover, arguably the considerable flexibility of the partial structures framework allows it to also accommodate non-Galilean infinite idealisations (Bueno and French, 2011). It is thus offered as a truly unitary approach to understanding the role and workings of idealisations—both Galilean and non-Galilean alike.

Is it enough for a realist to point to this meta-scientific framework as providing a satisfactory response to the challenges that idealisations pose to her? I do not think so. The framework of partial structures, partial homo-/isomorphisms, and quasi-truth allows us to identify a formal correspondence between an idealised model and its target, which in turn allows us to formally (re)present the idea that the model is in a sense 'latching onto' the target. Since idealised models can latch onto their targets in this sense, the framework thus 'accommodates' idealisations. But we should try to go beyond this by *accounting for* an idealised model's empirical success by showing how a model's 'latching onto' unobservable reality can be considered to be *responsible* for the model's predictive success. It is not clear how the existence of partial homo-/isomorphisms between (a formal representation of) a model and its target, or the model's quasi-truth for that matter, provides understanding of why the model is empirically successful *by virtue of* latching onto reality thus-and-so, and *regardless of* incorporating such-and-such aspects of misrepresentation. We should want a clearer sense of the role played by the idealising 'falsehoods' in an idealised model, and

[7] The realist faces the epistemic challenge of justifying her knowledge of the target-as-it-actually-is, of course, but this issue has nothing to do with idealisation per se.

a clearer sense of how the realist can bracket those aspects of the model as falling outside her realist commitments, despite them being useful, or even indispensable for making the predictions. The existence of partial homo-/isomorphism between an idealised model and a data model, for instance, says nothing about this in and of itself, and little has been said by way of analysing the explanatory credentials of such structural relations (vis-à-vis the idealisation challenge) in the context of the semantic view.[8]

* * *

Many have taken to heart the notion, well expressed by Giere, that when it comes to science 'idealisation and approximation are the essence, [so] an adequate theory of science must reflect this fact in its most basic concepts.' (Giere, 1988, p. 78) But in the face of the idealisation challenge 'reflecting' is not enough. An adequate (realist) theory of science should also *account* for the empirical success of idealised models, and the above accounts of idealisations fall short of throwing sufficient light on the roles played by idealisations in the production of predictive success. In particular, we should demand a clearer sense of how the realist can consider idealisations not to be the driving force behind models' predictive success, and how the realist can rather consider the models' latching onto reality to be responsible for it.

4 Towards a realist analysis of idealisations

Scientific models and their inexact representational fit to the world raise various questions, many of which specifically concern idealisations. But it is important to realise that the articulation-of-realism challenge, in particular, is actually *not specifically* about idealisations. Rather, it is an instance of a much broader challenge to realism. The general form of the question at stake is: how can a model that is false in *this* way be empirically successful in *that* way? This question arises in connection with any empirically successful model that incorporates falsehoods, regardless of the reason behind those falsehoods. A model can incorporate falsehoods due to being idealised, but also for other reasons. In particular, the same question arises even if scientists are simply mistaken or misguided about their target of theorising.

Recognising the general nature of the question at stake, it is immediately unclear why the realist response to it should vary depending on the reason behind the representational inaccuracy in play. Why would it matter for the realist response whether the reason behind a representational inaccuracy is

[8] It is possible that more can be said on behalf of the structuralist analysis of idealisation, and the partial structures analysis of idealisations can well be a useful part of a bigger picture, of course.

an intentional simplification (as in the case of idealisation), or an unintentional, erroneous assumption?[9] After all, in both cases the realist hopes to be able to answer this question in terms of how the model relates to its target, in such a way that we can regard the sense in which the model latches onto its the target as being responsible for the model's empirical success. Furthermore, if we have a fruitful conceptual framework for offering a realist response in connection with unintentional misrepresentations in science, it is reasonable to try to apply that framework also to idealisations (qua intentional misrepresentations).

My analysis of idealisations from a realist perspective is guided by this line of thought. That is, I adopt a conceptual framework that I have found fruitful and apposite in connection with some models that incorporate fundamentally misguided assumptions.[10] I claim that within this framework we can in quite general terms naturally account for idealisations' utility in modal terms, going beyond merely noting that idealisations traffic in non-actual possibilia to which actual systems can be usefully compared, or that there are different (quasi-formalisable) senses in which idealised models can be 'partially true' despite the 'falsehoods' they incorporate.

Here is an outline of the conceptual framework. We shall focus on predictive success of models, ignoring their explanatory success for now. (I will comment on the explanatory utility of idealisations later.) Given a particular model, I am interested in question Q: how is the model predictively successful—viz. empirically adequate in the relevant ways—despite misrepresenting its target in certain respects. To answer this question we can consider a range of models that vary in those respects, corresponding to a range of possible systems they can be taken to represent. The aim is to show how the required degree of empirical adequacy is independent of the particular false assumptions incorporated in the model. Independence is a matter of robustness of predictive success under variation in possible modelling assumptions that together with fixed background assumptions (including the relevant laws) yield the predictive success at stake.

The thought is that modal information of this sort can furnish a realist response to Q to the following extent: it shows how modelling assumptions can be false but nevertheless 'contain' veridical assumptions about the target that are responsible for the predictive success in the sense that variation in the specific false assumptions, without variation in the 'contained' veridical assumptions, would not undo the predictive success. It is via these

[9]There are questions about the modelling practice that specifically involve idealisations: for example, is the endemic and carefree employment of idealisations in tension with realism? My way of framing the idealisation-challenge focuses on models themselves, not the modelling practice.

[10]Some of these models have animated much discussion in the realism debate, such as Fresnel's elastic ether model of the partial refraction and reflection of light, used to derive the so-called Fresnel's equations. See Saatsi (2005).

'contained' veridical assumptions that the model can be viewed as latching onto reality so as to ensure the model's predictive success. The specific false assumptions involved, regardless of their indispensability or otherwise for presenting and working with the model, are not doing any of the heavy lifting in producing the predictive successes at stake.

The tricky business lies in spelling out the sense in which a set of modelling assumptions can 'contain' veridical, success-fuelling assumptions. (The complex literature on verisimilitude and approximate/partial truth demonstrates how difficult these issues are.) Here I take 'containing' to be a matter of the specific modelling assumptions together with the relevant background assumptions entailing some further, less specific features of the target system, such that getting these further features right (in conjunction with the relevant background assumptions, including laws) would suffice for a model to exhibit the predictive success at stake.[11] All this is perhaps best elaborated by illustrating it via a simple toy example. Before we get to this, I note again that nothing in the abstract outline above directly corresponds to the notion of idealisation. This is as it should be, for the reasons given at the start of this section.

As for a toy example, consider a model system with a graph-like structure. The model represents its target system as having four nodes, connected by some dyadic relations as in Figure 1.

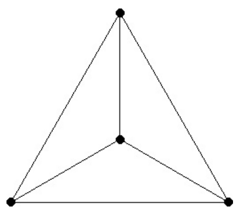

Figure 1: 3-regular model with four vertices.

That is, the model represents a target of four vertices, connected with one another in this 3-regular way. (A graph structure is 3-regular if each of the nodes is connected to three other nodes.) Assume that the relevant background assumptions, including the relevant laws, allow one to make a successful prediction about the system's behaviour under some circumstances (e.g. in the chemistry of carbon molecules.)

Assume further (for the sake of the argument) that the phenomenon in question, given the laws, is only exhibited by systems that have less than eight nodes, and that for such systems the phenomenon only depends on 3-regularity. That is, we are assuming that the relevant laws are such that

[11] By 'entailing' I mean not only logical entailment, but also metaphysical entailment, such as the relationship between determinate and corresponding determinable properties. If facts about such relationships can be packed into the background assumptions we can ensure logical entailment, of course.

even if the system were to have six nodes, say, and a 3-regular structure, it would display the behavior predicted.

Given these assumptions, representing correctly the number of nodes is clearly not relevant for the predictive success of our model. If as a matter of fact the target is 3-regular and has six nodes (as in Figure 2), then the model misrepresents the target regarding the number of nodes and relations between them, but it still 'gets right' the fact that each of the nodes is connected to three other nodes, i.e. the structure is 3-regular.

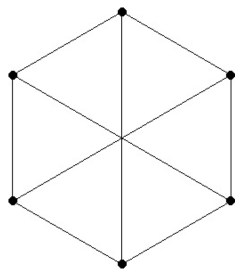

Figure 2: The actual 3-regular target with six vertices.

I take it that there is an intuitively clear sense in which our model gets the relevant feature of the target right: it latches onto reality by correctly representing the target's 3-regularity. This is the critical, less specific feature of the system that the model 'contains'. It is less specific than the modelling assumptions that specify which node is connected to which. (Note that 3-regularity need not be part of the stated modelling assumptions, and need only be 'contained' in these assumptions in the sense of being entailed by them.) And it can be this sense of 'containing' of the veridical assumption about the target—this sense of 'latching onto' the target—that explains the model's empirical adequacy vis-à-vis the phenomenon in question. The model's predictive success is explained in a way that renders wholly immaterial the misrepresentations the model incorporates with respect to the number of nodes, and which node is connected to which. In the setting of this toy-example, grasping this sense 'latching onto' the target adequately answers the challenge at stake.

For the very same reason a model that represents the target as a different 3-regular graph of six nodes (as in Figure 3) would be equally empirically successful. It also 'gets right' the fact that the target is 3-regular, and that it has less than eight nodes. As it happens, there are only three 3-regular graphs of less than eight nodes. A given target system can only instantiate one of these specific structures, but given the laws (we have assumed), the relevant features that our model needs to latch onto are less specific than that. The critical, less-specific features of the target are: the target has 4 or 6 nodes; the target is 3-regular. These less-specific modelling assumptions are realized in three different, more specific ways. Any model is going to incorporate one or another of the specific realisers, but all that really matters is that a model incorporates *one or another* of these features, i.e. that it incorporates the less specific feature. Some models can furthermore count as being idealised by virtue of incorporating such specific realiser that sufficiently simplifies the model in its presentation and operation. (In some context a 3-regular graph of mere 4 nodes could be an idealisation of a larger 3-regular graph, for example.)

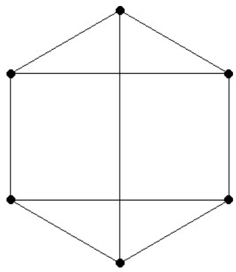

Figure 3: The other possible 3-connected model with six vertices.

This is merely a simple toy example, of course, but it serves to bring out the key features of an interesting conceptual framework. In particular, it shows how accounting for a model's predictive success can turn on grasping the robustness of predictive success under variation in the specific modelling assumptions that all 'contain' a critical veridical assumption.[12] The sense in which a model (in relation to the relevant background assumptions) can thus latch onto reality is conceptually quite straightforward, and not in my view well captured by the existing (quasi-)formal frameworks for 'partial truth', approximate truth, or verisimilitude. What matters is the grasping of what is common to different possible systems, such that the common feature is all that matters, since variation in other features is immaterial: any model that features some 'realiser' of the common feature would count as predictively successful. A derivation of the prediction further requires the right laws of nature as background assumptions, grounding the relevance of these less-specific features.

The sense in which modelling assumptions can 'contain' a veridical, success-fuelling assumption need not be captured by a notion of partial truth applied to propositions that can be used to specify the model. Consider, for example, the model:

> {Alice knows Bob. Bob knows Erik and Fred. Erik knows David and Charlie. David knows Fred and Alice. Fred knows Charlie. Charlie knows Alice.}

[12] This has connotations of robustness analysis of idealised models. (See e.g. Odenbaugh, 2011) Exploring the connections to the literature on robustness analysis requires further work. (Thanks to Arnon Levy for flagging this question for me.)

This model can latch onto the target represented by

> {Alice knows Bob. Bob knows Erik and Charlie. Erik knows David and Fred. David knows Alice and Fred. Fred knows Charlie. Charlie knows Alice.}

The two systems represented by these two sets of propositions exhibit the two alternative 3-regular structures with six nodes. Neither set of propositions explicitly says anything about the shared 3-regularity, however, and the underlying similarity is not explicitly represented by the propositions, nor revealed by looking at the (partial) truth or otherwise of the (sets of) propositions involved in presenting the two systems. Since the pertinent similarity between the model and the target need not be part of the explicit representational content of the model—the model need not represent the target *as* 3-regular—I call the model *inferentially* veridical (as opposed to representationally veridical). The idea is that from the model we can infer, with the help of the relevant background assumptions, the critical veridical assumptions.[13]

One may worry that this kind of 'inferential veridicality' is too thin to support a realist account of empirical success. One may worry, for example, how the less specific feature 'having 4 or 6 nodes'—a disjunctive property—can be attributed to the target. Or one may worry about the sense in which a model 'containing' a veridical assumption of this kind can account for the model's empirical adequacy in a realist spirit. I think the right response to such worries is to note that it is the appeal to laws of nature in deriving predictions from a model that underwrites the significance of the less-specific properties, regardless of whether or not they have disjunctive realisations. So, given these laws, from a scientific point of view such a property can be a genuine, bona fide feature of the world on which our theorising can latch, despite its disjunctive (or unspecific, or vague) character. One way to put this is to say that with the less-specific, veridical assumptions we are latching onto an important modal truth: had the target had only 4 (as opposed to 6) nodes, all with 3 connections, the same result would have ensued given the relevant laws of nature.

One may push the same worry in more general terms, in relation to my characterisation of how the veridical assumptions are 'contained' in the model. I said above that 'containing' is a matter of the specific modelling assumptions together with the relevant background assumptions *entailing* some further, less specific features of the target system. The worry here is that this idea that the model is thus latching onto some less specific, more abstract worldly features seems to face a 'disjunction problem': since any

[13] The realist can then claim that derivations of successful predictions involve such inferences, and thus involve the veridical assumptions. Cf. Saatsi (2005) for related discussion in connection with Fresnel's model of light.

modelling assumption p always entails $p \vee q$, any model is (allegedly) guaranteed to latch onto reality, as long as there is *some* true q such that it would work to produce the right prediction.[14] Does a model's inferential veridicality thereby become a trivial matter, deflating realism of any worthwhile commitment?

The answer is no. It is *not* the case that any model is guaranteed to latch onto reality just by virtue of being predictively successful, since a model latches onto reality partly by virtue of appealing to appropriate facts about laws of nature. For example, if one constructs an empirically adequate model M in classical Newtonian physics of a purely quantum phenomenon, the false modelling assumptions are not latching onto reality, since there is no possible classical model that provides a faithful, veridical representation of the target. It is not the case that some more complicated classical model faithfully represents the system and shares the critical, less specific properties with M such that any classical model that exhibits those properties would be equally empirically adequate as M. For the same reason a Ptolemaic model with epicycles does not latch onto its target (the solar system) despite its impressive empirical success.

Admittedly there is much more to be said regarding the kind of realism that can be served by the conceptual framework I am proposing here, and I hasten to add that it is not the case that realist intuitions and cause are saved *just* by showing predictively successful models being inferentially veridical. There can be interesting cases of local underdetermination where radically different modelling assumptions, in conjunction with the right laws, give rise to more or less the same predictions. (See for example Saatsi and Vickers (2011) for one such case.) In such cases the explanation of predictive success can have a strong anti-realist flavour. But in many cases the details of the derivation, and in particular the role played therein by the relevant less-specific features (with respect to which the model is inferentially veridical), can serve the realist cause by saving the 'no miracles' intuition. Or so I contend.

5 Beyond toy examples

I have proposed, largely in the abstract, a conceptual framework for accounting for the predictive success of idealised models in modal terms. One may wonder whether this conceptual framework can capture some real idealised models as well. I certainly think so! Consider a paradigmatic Galilean idealisation, such as an ideal pendulum as a model of my grand father's pendulum clock. The model's degree of empirical adequacy is naturally accounted for in terms of the model's inferential veridicality, in conjunction

[14] See Strevens's (2008) discussion of the disjunction problem in connection with his difference-making account of causal explanation that operates by abstraction.

with the appropriateness of the background laws (Newtonian mechanics + gravity). The model is inferentially veridical by virtue of entailing truths about less specific features of the target such that any model that realises those features in one way or another will attain at least that degree of empirical adequacy. The relevant less specific features concern a vague force function, vague specification of the pendulum's dimensions, etc. The ideal pendulum model represents a particularly simple specific realisation of these less specific (vague) features, and its empirical adequacy is easily accountable—regardless of its misrepresentation in these respects—by noting the robustness of its predictive success under variation in the particular false specification of the critical less specific features, the specification that constitute the idealisation.

Various other Galilean idealisations similarly lend themselves to analysis in these modal terms. (See Saatsi (2011) for further discussion.) One might wonder how much we gain from this, given that arguably Galilean idealisations do not present a serious challenge to realism to begin with. Although I already admitted (§2) that realists have a wealth of resources in responding to a challenge posed by Galilean idealisations, I think the realist can further gain from the conceptual framework advocated here. In particular, the framework allows us to shed further light on the modal aspects of idealised models, and how those aspects can feed into an account of an idealised model's empirical success. This framework affords us a better sense of a particular way in which an idealised model can latch onto reality so as to account for the model's empirical success.

Furthermore, there are reasons to think that the framework can also deal with (at least some) non-Galilean idealisations. The distinction between Galilean and non-Galilean idealisations need not be as deep as one might think. In relation to the much discussed infinite continuum idealisations in statistical physics, for example, we may construe the distinction in terms of how indispensable a given idealisation is to a model. On one side we have Galilean idealisations which are *controllable*, at least in principle, in the sense that we can we can replace our original model with a related, less idealised model that represents the system in question more truthfully (for example by including previously omitted forces). On the other side we have *uncontrollable*, non-Galilean idealisations that cannot be thus eliminated or reduced, even in principle, by a related, less idealised model. A paradigmatic example of such uncontrollable idealisation is the use of the thermodynamic limit in statistical physics of finite systems, where the number of particles n and the volume V of a system are taken to infinity while keeping n/V constant. This mathematical idealisation is uncontrollable since it cannot be replaced with a model that takes n to be some finite-but-large number (e.g. $\sim 10^{23}$), thereby representing better the finitude and atomicity of the actual system.

The sense of indispensability of such uncontrollable idealisations raises interesting questions, and it clearly in some sense demarcates these idealisations from the controllable cases. The uncontrollability in and of itself does not mean that these models cannot be viewed as inferentially veridical, however. What it means, rather, is that we are unable to construct models that are more veridical in these idealising respects, so as to demonstrate *in that way* how the predictive success of the idealised model is robust under variation in the idealising assumptions. Models incorporating uncontrollable idealisations can still be inferentially veridical, however, in the sense that it can be a modal fact about the relevant laws of nature that they deductively yield, when combined with more veridical assumptions about the idealised features, the same or improved degree of empirical adequacy. *Our* (in)ability to demonstrate this—in principle or in practice—by de-idealising the original model need not necessarily be taken to indicate that such fact does not obtain.

There is a close analogy here with debates concerning mathematics' indispensability to science. Nominalists argue that regardless of our inability to nominalize our best theories we can maintain that it is the non-mathematical content of our theories that is responsible for the theories' empirical success, with mathematics playing a role only in representing non-mathematical facts and facilitating reasoning about it.[15] In a similar spirit I maintain that the indispensability of the uncontrollable infinite limits in statistical mechanics, for instance, can be indispensable only for representing and reasoning about systems with enormous but nevertheless finite numbers of components. It can still be a modal fact about the relevant micro-level laws of nature that they entail the same empirical results from veridical assumptions about the interacting micro-constituents.

But how, one may wonder, can this attitude be justified, if not by having good reasons to think that a model is de-idealisable, at least in principle? The answer is that one's understanding of the workings of an uncontrollable idealisation can involve much else besides the assumptions that go into a particular non-Galilean model. That is, the full set of theoretical resources that can come to bear on justifying one's belief in such modal fact about the laws—justifying the inferential veridicality of the idealised model—goes well beyond the modelling assumptions. In the full theoretical context of such models we can arguably explain, by reference to relevant facts about finite systems, why an infinite mathematical idealisation is empirically adequate to the degree it is, notwithstanding its indispensability. This broader theoretical contexts has been extensively discussed in the recent literature. (See e.g. Butterfield, 2011a,b; Menon and Callender, 2013; Norton, 2012) It is through such theoretical accounts of a given uncontrollable idealisation that we get a handle on the sense in which the model 'gets right' some critical less specific features of large-enough systems. These are the features

[15]See e,g. Melia (2000) and Saatsi (forthcoming).

that the model shares with large finite systems, features that in conjunction with the relevant laws entail the right predictions (to a sufficient degree of approximation).

The details of these 'reductionist accounts' of the continuum limit in statistical physics remain to be discussed further in the context of my conceptual framework. I have to leave this for further work, and move on to conclude the paper with brief remarks on explanation. Throughout the paper I have focused on the predictive success of idealised models, largely bracketing the role of idealisations in successful scientific explanations. The explanatory dimension also matters to the realist, of course, given the role of inference to the best explanation in many realist gambits, for example. (It is worth noting that Batterman's much discussed work on uncontrollable idealisations almost exclusively concern their explanatory indispensability.) It is impossible for me to do justice to this rather large topic here, but let me just note the importance of considering models' explanatory successes quite separately from their predictive successes. The distinction between predictive and explanatory success was perhaps only of minor consequence back in the day of the DN-model of explanation. But in the contemporary context, largely ruled by different modal accounts of explanation, the conceptual difference between prediction and explanation matters a great deal to the way realists should apportion their epistemological commitments in relation to scientifically successful theories and models. Different issues come to the fore in accounting for the explanatory role played by the falsehoods that constitute idealisations. The indispensability of idealisations for explanations, for example, raises issues for the realist that are closely related, or analogous to the issues raised by the arguably indispensable role that mathematics plays in scientific explanations. I have argued elsewhere that the realist should consider the latter issues in close contact with well-formed views about the nature explanation (Saatsi, forthcoming). I believe the same holds for the former issues as well.

References

Batterman, R. W. (2005). Critical phenomena and breaking drops: Infinite idealizations in physics. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics 36*(2), 225–244.

Butterfield, J. (2011a). Emergence, Reduction and Supervenience: A Varied Landscape. *Foundations of Physics 41*(6), 920–959.

Butterfield, J. (2011b). Less is Different: Emergence and Reduction Reconciled. *Foundations of Physics 41*(6), 1065–1135.

Da Costa, N. and S. French (2003). *Science and partial truth: a unitary* *approach to models and scientific reasoning*. Oxford University Press, USA.

Giere, R. (1988). *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press.

Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell.

McMullin, E. (1985). Galilean idealization. *Studies in History and Philosophy of Science 16*, 247–273.

Melia, J. (2000). Weaseling Away the Indispensability Argument. *Mind 109*, 455–479.

Menon, T. and C. Callender (2013). Ch-Ch-Changes Philosophical Questions Raised by Phase Transitions. In Robert Batterman (Ed.), *The Oxford Handbook of Philosophy of Physics*, pp. 189. OUP Usa.

Niiniluoto, I. (2007). Idealization, Counterfactuals, and Truthlikeness. In Jerzy Brzeziski, Andrzej Klawiter, Theo A. F. Kuipers, Krzysztof Lastowski, Katarzyna Paprzycka, and Piotr Przybysz (Eds.), *The Courage of Doing Philosophy: Essays Dedicated to Leszek Nowak*, pp. 103–122. Rodopi.

Norton, J. (2012). Approximation and idealization: Why the difference matters. *Philosophy of Science 79*(2), 207–232.

Odenbaugh, J. (2011). True Lies: Realism, Robustness and Models. *Philosophy of Science 78*(5), 1177–1188.

Saatsi, J. (2011a). Idealized models as inferentially veridical representations: a conceptual framework. In P. Humphreys and C. Imbert (Eds.), *Representations, Models, and Simulations*, pp. 234–249. London: Routledge.

Saatsi, J. (2011b). The Enhanced Indispensability Argument: Representational versus Explanatory Role of Mathematics in Science. *The British Journal for the Philosophy of Science 62*(1), 143–154.

Saatsi, J. (forthcoming). On the 'Indispensable Explanatory Role' of Mathematics. *Mind*.

Saatsi, J. and P. Vickers (2011). Miraculous success? Inconsistency and untruth in Kirchhoff's diffraction theory. *The British Journal for the Philosophy of Science 62*(1), 29–46.

Saatsi, J. T. (2005). Reconsidering the Fresnel-Maxwell case study. *Studies in History and Philosophy of Science 36*, 509–538.

Sorensen, R. (2012). Veridical idealizations. In M. Frappier, L. Meynell, and J. R. Brown (Eds.), *Thought Experiments in Philosophy, Science, and the Arts*, pp. 30–50. Routledge.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation.* Harvard: Harvard University Press.

Toon, A. (2012). *Models as Make-Believe: Imagination, Fiction, and Scientific Representation.* Palgrave Macmillan.