Semantic Realism in the Semantic Conception of Theories

January 16, 2020

Abstract

Semantic realism can be characterised as the idea that scientific theories are truth-bearers, and that they are true or false in virtue of the world. This notion is often assumed, but rarely discussed in the literature. I examine how it fares in the context of the semantic view of theories and in connection with the literature on scientific representation. Making sense of semantic realism requires specifying the conditions of application of theoretical models, even for models that are not actually used, which leads to several difficulties. My conclusion is that semantic realism is far more demanding than one would expect. Finally, I briefly examine some pragmatist alternatives.

1 Introduction

Scientific realism is often presented as the combination of three theses: metaphysical realism, semantic realism and epistemic realism (Psillos, 1999). Epistemic realism has by far been the main focus of philosophical discussions in the recent decades, and semantic realism is generally uncritically assumed, if mentioned at all, in the form of vague slogans such as "scientific theories should be interpreted literally", or "at face value". The object of this article is to examine this notion, and in particular how it fares in relation to the so-called semantic conception of theories, according to which scientific theories are best presented as collections of models.

Semantic realism can be characterised as the combination of a truth-conditional semantics and a conception of truth that is not epistemically constrained (Shalkowski, 1995) (or such that truth-conditions are "potentially evidence transcendent" (Miller, 2003)). This is generally summarised in the philosophy of science literature as the idea that scientific theories are truth-bearers and that they are true or false *in virtue of reality*. The first aspect distinguishes semantic realism from what Psillos (1999) calls "eliminative instrumentalism", which takes theories to be mere instruments without truth-values, and the second aspect distinguishes it from "reductive empiricism" (among other positions, arguably), which attempts to reinterpret the content of scientific theories in terms of mere observables.

For a semantic realist, scientific theories describe a mind-independent reality, and therefore their content should not be interpreted in terms of notions such as measurement, observations, intentions, information, social norms or any other epistemically loaded or anthropocentric term, at least not if these are taken to be unanalysable, irreducible notions associated with the users of the theory. This idea seems to be an important desideratum in the metaphysics of science (Bell (2004)'s disdain of the notion of measurement is often cited to that effect in the philosophy of physics literature). Not all positions called "realism" satisfy this condition. For example, Putnam's internal realism explicitly rejects semantic realism. Nevertheless, it is generally accepted as an essential component of scientific realism.

The semantic conception of theories has now become common-place in philosophy of science (Suppe, 1989; Suppes, 1960; van Fraassen, 1980; Ladyman and Ross, 2007). According to this conception, scientific theories are not sets of statements about the world. They are best presented as collections of models. Theoretical models are abstract entities that can be used to represent target systems. They can be understood either as specific mathematical structures, for example state-spaces and transitions between states, or, following the Tarskian tradition, as set-theoretical structures that, once mapped to a vocabulary, satisfy theoretical statements. The general idea of the semantic conception is that the theoretical statements found in scientific textbooks do not constitute the theory, but merely describe, in a particular language, the families of theoretical models that do constitute it (the semantic view "construes theories as what their formulations refer to when the formulations are given a (formal) semantic interpretation" (Suppe, 1989, p. 4)). There can be more than one way of describing these models, and more than one formulation of a theory. The view was developed in reaction to problems affecting the statement view entertained by logical empiricists, and purports to be more connected to actual scientific practice (Lutz, 2015).

There is a straightforward tension between semantic realism and the semantic conception of theories, insofar as one of the main purposes of the latter was to acknowledge that scientific representation is not linguistic and thus to get rid of problematic issues falling under the scope of philosophy of language. A model, contrarily to a linguistic statement, is not generally said to be true or false: instead it is said to be good or bad, or accurate or inaccurate. How, then, shall we understand this idea that theories are truth-bearers?

This problem was acknowledged by van Fraassen and Ransanz (1985), Giere (1991, p. 85) and Suppe (1989), who claimed that a scientific theory should also be qualified by one or several statements asserting something about how the models of the theory relate to reality (for Giere, "various hypotheses linking those models with systems in the real world", for Suppe, "a theoretical hypothesis claiming that real-world phenomena [...] stand in some mapping relationship to the theory structure"). Presumably, this will hinge on a notion, call it veridicality, that is the counterpart of truth when applied to models. A veridical model is a faithful representation of its target. A natural candidate for veridicality would be the idea that the model corresponds to its target, or that there is an isomorphism between the model and the (causal, modal or only extensional) structure of the target. But introducing a notion of veridicality can only be the first step. We should also explain the link between this notion, which applies to models, and truth at the theory level: what kind of statement do we need to qualify the models of the theory? For instance, do we want all models to be veridical, or only some of them? This is the question that this article will address.

My aim is to show that it is much more complex than one could think at first sight, in light of recent discussions on the topic of scientific representation. I think that the main conclusion to draw from my analysis is that semantic realism is a demanding notion, and that we should not assume it uncritically. Another conclusion is that the weakened versions of scientific realism that have been proposed in recent decades to accommodate anti-realist arguments, such as structural realism, are hardly compatible with it: I will argue that these positions cannot be only *about the world* if they eschew reference to natural properties. In conclusion, I briefly examine pragmatist alternatives to semantic realism.

But let me first say a word about the notion of veridicality.

2 Veridicality

Let me begin by examining how we shall understand the notion of veridicality, which purports to be the counterpart of truth for theoretical models. We have two main options at our disposal to define it, but I think they more or less boil down to the same idea.

The first option is to define veridicality directly as a relation between the model and what it represents; for example, an isomorphism, a partial isomorphism (Bueno and French, 2011; van Fraassen, 1980), or something less formal like "similarity" (Giere, 1991).

The second option is to make a detour through language, by considering models, qua set-theoretical structures, as truth-makers for statements, in the Tarskian tradition. These statements can include theoretical laws and principles as well as specific assumptions about the target of representation. Considering that the targets of models can also act as truth-makers for theoretical statements, we could say that a model is veridical if all the statements it satisfies are also satisfied by its target.

The first option involving a direct relation between models and their targets should be qualified, and it is reasonable to assume, as many commentators have observed, that language must play a mediation role at some point in order to fix the intended target of representation. For example, someone assuming that the model-target relationship is some kind of isomorphism should provide a set of "important" objects, properties and relations, the structure of which is preserved by isomorphism, to avoid triviality (Ainsworth, 2009). As for similarity, it can be argued that making explicit in which respects two objects are similar also requires linguistic mediation (Chakravartty, 2001). In general, language is needed to say what the theory is about (Frigg, 2006; French and Saatsi, 2006; Thomson-Jones, 2012), and relate the various models of the theory together (Halvorson, 2012). This means that veridicality only makes sense given a certain interpretation of the model in terms of the target and that presumably, the way models are interpreted should be consistent across various uses of a given theory.

If we accept this role for language, the two options mentioned above can be considered roughly equivalent, given that the notion of truth involved in the linguistic option is under-specified, and could be specified in such a way that it corresponds to one's preferred target-model relationship (one might wonder whether the model-target relationship grounds linguistic truth, or the other way around, but I won't address this question–see Chakravartty (2001) for discussion). So the linguistic formulation is more general, and it is the one I will adopt, although not much hinges on that. Let me remark that this formulation assumes that models are not bare set-theoretical structures, but structures mapped to a theoretical vocabulary (such that "important" objects, properties and relations in the domain of object are "given a name"). This is required for models to act as truth-makers. But given the previous remarks on mediation by language, this assumption seems reasonable.

3 From Model Veridicality to Theoretical Truth

Let us now examine the possible routes from veridicality to truth. What statement should qualify a scientific theory, once we have presented its models?

A first option would be this: a theory is true if and only if all its models are veridical. Obviously, this will not do for several reasons, the main one being that it makes no sense to talk about veridicality without specifying which target the model represents. Another reason is that arguably, some (if not most) models of any theory do not purport to represent any concrete target, so they cannot really be veridical. This is true at least considering that all structures that satisfy the laws and principles of the theory are models of this theory (taking theories to be presented by means of sets of laws and principles, following the model of physics), but even considering a different notion of theoretical models, one can easily come up with examples of scientific models of a theory that do not purport to have any real target in the world, for example, in thought experiments.

A second option would be: a theory is true if and only if all models that represent a target in the world are veridical.

At this point, we should say more about what it is to represent. There are various accounts of representation. It has been argued that models represent their targets in virtue of a relation of isomorphism or partial isomorphism (van Fraassen, 1980; French and Ladyman, 1999; C. A. Da Costa and French, 2003) or similarity (Giere, 1991). If this echoes the discussion about veridicality in the previous section, this is no coincidence: these naturalistic accounts have been criticised precisely for not distinguishing between representation and accurate representation, and not accounting for the possibility of misrepresenting a target (Suárez, 2003) (but see Bueno and French, 2011).

The last point about misrepresentation is certainly important for our purpose: the semantic realist *needs* this distinction between representation and accurate representation, otherwise the idea that theories are true about the phenomena they can be used to represent becomes trivial, because representing and being true would amount to the same thing. We would not want to say that it is impossible to represent the solar system in Newtonian mechanics, for example, only because Newtonian models are not perfectly accurate. To be either true or false in virtue of the world, a theory must be capable of having non-veridical models that still represent their targets.

The common strategy to account for misrepresentation is to understand representation as (at least) a three-place predicate between a vehicle, a target and a user: the user takes the vehicle to stand for, or *denote*, the target, or in Suárez (2004)'s terms, it has *representational force*, and although norms of representation could be involved in the process, this says nothing about the model's accuracy. It has now become commonplace in the literature on scientific representation to assume that the user indeed plays a role (van Fraassen, 2008; Giere, 2010; Contessa, 2007).

However, assuming this role for the user, it appears that our second option for defining theoretical truth, which is restricted to the models that represent a target, is too limited, because we want theoretical truth to extend beyond actual representational use. For example, we would like to say that Newtonian mechanics is false if a phenomenon does not respect its predictions, even if no scientist has ever attempted to model this phenomenon (or that it was *already false* before astronomers observed that the trajectory of Mercury did not conform to its predictions). This is required for a semantic realist: theoretical truth must not be tied to particular uses of the theory, since it must not be epistemically constrained. So we must specify conditions of application even for models that are not actually used (this problem does not affect naturalistic accounts of representation that dispense with a user in the same way, but they also need to specify conditions of application if they want to account for misrepresentation, so the rest of the discussion will be relevant to them).

In order to circumvent this problem, let me introduce the notion of applicability, which corresponds to the idea that a model *could* be used to represent a target. By this, I mean that it respects certain norms of representation imposed by the theory. These norms presumably encapsulate Giere and Suppe's hypotheses that would provide a "link" or "mapping" between the theory and the world, or at least part of them: they tell us what a given model could potentially represent in the world, and how, without implying that the model is veridical. The modal aspect is important: I do not mean that the model is actually used to represent the target, only that it would be legitimate to use it. This makes a distinction between applicability and representational force, at least for accounts of representation that involve a user and a denotational aspect.

It could be objected that whether or not a model is legitimate to represent a target depends on a context of use, on particular users and associated aims, and on contextually informed practical knowledge. A model could be appropriate to *predict* a measurement outcome while experimenting on a target, but inappropriate to *explain* some aspect of the same target. But the fact that different models are more or less appropriate for different uses is compatible with the idea that general, a-contextual norms apply to all possible contexts of use, and we could say that it is enough for a model to satisfy these general norms to be applicable, and to be either veridical or not. Each legitimate application would provide an interpretation of the model in terms of the target, from which veridicality could be assessed. Perhaps this will not do, because other contextual factors are implied in the assessement of veridicality, or perhaps there are no a-contextual norms of representation. But the semantic realist should deny this, otherwise theoretical truth will be epistemically constrained, in that it will depend on what and how we choose to represent. The semantic realist needs an a-contextual notion of applicability, from which veridicality can be assessed.

Let us say, then, that a model applies to a target if it could legitimately be used to represent the target, even if it is not actually used, that is, if it respects a-contextual norms of representation. Note that these norms could be quite liberal. I noted in the previous section that veridicality can only be assessed if the model is interpreted in terms of the target. At this point, I do not want to assume that there is a univocal way of doing this; there might be several interpretations available for the same model and target, each corresponding to a potential application of a given model. I will say more on these aspects in the next section.

With the notion of applicability in hand, we could say that a theory is true if and only if all applicable models are veridical. But it is easy to find examples of theories that are not taken to be false, even if some models of the theory are not taken to be veridical of the target they represent: typically, idealised models do not represent their targets accurately; they caricature their targets (examples of this are frictionless planes, infinite gases or point masses). Now this could be taken to imply that these theories are actually false (Cartwright, 1983). But this option might not be very satisfying for a semantic realist, at least if she wants theoretical truth to matter somehow: idealisations are not generally taken to be a defect of representation, but rather a feature that contributes to the goodness of a scientific model. Besides, one could argue, contra Cartwright, that idealised models are mere placeholders for veridical models. The idea would be that idealisations can be "de-idealised" in principle, or that they would foster cognitive virtues, such as simplicity and computability, while retaining a veridical component concerning the relevant variables (McMullin, 1985; Elgin and Sober, 2002; Saatsi, 2016; Bokulich, 2016). We could expect these veridical components to figure in a non-idealised model of the theory.

In any case, idealisations are not the only example to illustrate this problem with our definition. Another example is when a model proposed to account for a phenomenon turns out to be inaccurate, and is later replaced by a better one within the same theory: the first model was not veridical, but the theory cannot be considered false for this reason, because it can provide a better model. So we need refinement.

Taking into account the idea that idealised models are placeholders for veridical ones, let us propose a final option: a theory is true if and only if for any potential target in the world, there is at least one model in the theory and one interpretation for which the model is a veridical representation of the target. This would be, in the case of idealisations, the "de-idealised" model (assuming that the theory is true). The idea is simply that a true theory *could* represent anything veridically, while a false theory could not.

For ease of comprehension, let me formalise this definition, taking $A(M, t)$ to be the set of potential legitimate applications of a model to a target (or interpretations of the model in terms of the target) and $V(M, a, t)$ the relation of veridicality given a certain application a. This gives:

$$Truth(T) : (\forall t)(\exists M \in T)(\exists a \in A(M, t))V(M, a, t)$$

This looks like a good option. We could refine it a bit, so as not to include potential targets that are not in the domain of application of the theory: we cannot blame the theory of evolution for not accounting for electrons. This could be formalised in terms of the existence of at least one model-application pair for the target. But I will omit this for the sake of simplicity.

There is a first difficulty with this account: the models we would like to be veridical, that is, the non-idealised models, might well be beyond even potential use. For example, most theories in physics do not have analytic solutions beyond very simple cases. This means that the corresponding models cannot be known by finite means. This does not mean that they could not be used if they could be known, but it is hard to imagine how we could handle a mathematical structure that cannot be described by analytic functions, and how we could interpret it in terms of a target. The notion of applying ends up being quite metaphysical in spirit if we want it to be extendable to this kind of model, and the notion of an epistemic user becomes quite idealistic.

But this difficulty is unavoidable for the semantic realist. Refusing this extension would threaten to render all theories false. So I think this definition of theoretical truth is the best option at our disposal.

I will address other complications with the notion of applicability in the next section and talk more about the notion of interpretation. For now, let us consider one last option for defining theoretical truth, which is van Fraassen's preferred option (van Fraassen and Ransanz, 1985). In substance, it says that a theory is true if it has at least one model that is veridical of the universe.

I think this option is problematic for different reasons. For example, it is not clear that the universe as a whole can be considered a legitimate target of veridical representation, if only because the users and vehicle of the representation are contained in it. It also exacerbates the problem of models that could only be used with infinite cognitive capacities. Another problem is that the cognitive content of the theory is reduced to only one of its models, and it becomes puzzling why theories should be collections of models in the first place. This idea is rather disconnected from actual scientific practice.

In any case, assuming reasonable principles, van Fraassen's solution does not have very different implications from the one that has just been discussed. If a model of the universe of a given theory can be separated into parts that are also models of the same theory, we can expect any potential target of representation for this theory to be represented by one of these parts, and if the model is veridical, we can expect its parts to be veridical as well. This means that for any target in the domain of application of the theory, there is a veridical model of it, and we are back to our former definition. These expectations are reasonable because they seem required for concrete representational uses: we need to represent specific, bounded targets; this is the primary use of theories. And scientists generally know how to split models into parts or how to combine them into larger models.

So I think it is better to assume the previous option, simply because it is weaker and more connected to actual representational uses. However, this choice will not deeply affect the arguments of the next sections.

4 The Problem of Conditions of Application

So far, we have introduced a notion of model veridicality, and we have defined theoretical truth in terms of the veridicality of at least one applicable model of the theory for any target. Now come the difficulties I was mentioning. They all have to do with the notion of applicability.

The main point, with this notion, is that we need something to connect the various models of a theory to their potential targets for the notion of theoretical truth to make sense at all. This is the role of applicability. But how to do so is not obvious.

I defined applicability as potential representational force, which is that by which a model represents a target in particular uses, so let us examine the latter notion. (I will restrict my presentation to inferentialist accounts of scientific representation, assuming that the rest of the discussion applies to more substantial, including naturalistic accounts. There is no reason why the same problems would not affect them.)

According to Suárez (2004, p. 768), representational force is "the capacity of a source to lead a competent and informed user to a consideration of the target". This includes having a denotational function (Suárez, 2015a, p. 44). The notion of denotational function is meant to account for representation of fictitious entities, but in the case of a concrete target, this means simple denotation: the model "points to" the target, which can be achieved by mere stipulation. Obviously, this idea is tied to particular uses and should not appear in the notion of applicability (at most we could retain the idea that targets must be accessible to users in principle, so as to be potentially denoted–this will be discussed later).

Beyond denotation, a scientific representation also allows users to make valid inferences on the target. This is what distinguishes epistemic (including scientific) representation from mere symbolic representation. These inferences need not be sound, in which case the model misrepresents its target.

According to Suárez, these conditions are not sufficient for being a representation. Other conditions might be involved in specific contexts. He claims, in particular, that there are norms of valid inferences within some representational practice: inferences must be licensed by epistemic community (Suárez, 2015b; Boesch, 2017). But according to him, nothing more can be said about representation in full generality.

Contessa (2007)'s account purports to be more substantial. It is cast in terms of interpretation. Interpreting a model means taking specific properties, relations or functions of the model to denote properties, relations and functions of the target. In brief, this means providing a mapping between relevant parts of the model and target, which is exactly what we meant by "interpretation" so far. This, according to Contessa, is enough to fulfil Suárez's conditions of representation (because it allows the user to make inferences on the target), and no more conditions are required. In particular, Contessa assumes that any mapping will do. His account is therefore more substantial, but also more liberal than Suàrez's.

For a semantic realist, the idea that no fully general account of representational force (hence of applicability) can be given is unsatisfactory. There is again a risk of running into problems of epistemic relativity. There should be, at least, a general account at the level of theories (maybe each theory or paradigm has its own). Contessa's solution is interesting for a semantic realist, because it is substantial, general, and also minimalist. Furthermore, it provides a formalisation of our notion of interpretation. So let us assume this account, including its notion of interpretation, and say that a model is applicable to a target if there exists an interpretation of the model in terms of the target.

What I wish to discuss now is whether one should assume (contra Contessa) particular constraints on the interpretations that are legitimate for semantic realism to make sense: shall we allow any mapping between the objects, properties, relations and functions of the model and the target to count as a possible application? To take Contessa's example, can we use Rutherford's model of the atom to represent a hockey puck sliding on ice, taking the nucleus to denote the ice and the electron to denote the puck?

Maybe Contessa's liberal attitude is warranted if we are interested in what representation is in general, but in the context of assessing what it means for theories to be true or false, we need more constraints. There are at least two criteria to be met for our definition of theoretical truth to make sense. The first one is that these constraints shouldn't be too strict, so as not to make applicability equivalent to veridicality, otherwise any theory will end up being true, insofar as when its models are non-veridical, it is simply because they do not apply to the target. The second criterion is that these constraints shouldn't be too loose, otherwise almost any theory will have a model that is veridical for any target, even when we would intuitively consider this theory false.

The first problem is related to the idea mentioned above that a good account of representation should also account for misrepresentation: a model can represent its target inaccurately. To illustrate this problem, imagine that we have a theory of water that says that it is composed of H_2O molecules. As semantic realists, we want this theory to have truth-conditions. We want to say that the theory is true if, for any potential target of the theory, that is, any amount of water, there is one model of the theory that will be veridical, a model with H_2O molecules suitably configured. This first problem would occur if we were to define water (that is, the domain of application of the theory) circularly, as the kind of substance that is composed of H_2O molecules. Then our theory would be trivially true (note that this problem affects van Fraassen's version of theoretical truth as well, assuming that a domain of application for the theory must be specified: this domain would be limited to the phenomena in the universe for which the theory is accurate).

As for the second problem, we can illustrate it with Contessa's account of representation. Recall that according to Contessa, any interpretation of the model in terms of the target is allowed. Going from actual use to potential use, this means that any model applies to any target whatsoever, and in many different ways. This liberal conception of representation has been criticised by Bolinska (2013), who claims that more is required for a vehicle to represent a target system. In particular, the agent must aim to faithfully represent the target system using that vehicle, and the vehicle must be informative about the target. An arbitrary mapping does not satisfy these conditions. Contessa's liberalism is particularly problematic for a semantic realist attempting to define theoretical truth. It seems too easy for a theory to be true in this context: can't we just find a mapping, however contrived, between our target and a model, however ad-hoc, that will be such that the model is veridical? One can take the C in CO_2 to denote oxygen, and the O to denote hydrogen, and the theory that water is CO_2 will be vindicated. But then truth does not depend on the world any more. This is reminiscent of Putnam (1980)'s model-theoretical argument against metaphysical realism, and of Newman's objection against structural realism (Demopoulos and Friedman, 1985).

So there must be restrictions on which models apply to which targets, and how: not just any mapping will do.

Intuitively, one could think of these restrictions as conceptual norms governing the interpretation of theoretical vocabulary, applying across various contexts of use. The idea would be that the vocabulary involved in the model-target mapping should be interpreted consistently in terms of target properties. There is still room for interpretation given these constraints: specific models can employ a vocabulary that is not present at the theory level, hence not constrained in the same way (proper names as opposed to predicates for example). In any case, such constraints do not necessarily entail a collapse of applicability into veridicality.

I believe that a semantic realist must assume something along these lines for her position to make sense. Importantly, the idea cannot be that the vocabulary involved is interpreted only in terms of the structure of the theory: what is needed is to fix the *denotation* of this vocabulary, or its extension in the world, so as to assess that the structure of the theory gets things right. Assuming that the vocabulary involved in conditions of application is defined only by the theoretical structure would amount to saying that the theory applies only when its structure gets things right, and as we have seen, this is incompatible with semantic realism.

Does this idea of consistently interpreting the vocabulary solve all our worries? I think we still face a dilemma. Assume that theoretical predicates are interpreted in such a way that they consistently refer to the same properties or relations in all uses of the theory. Should we also add that particular objects in the model must denote particular objects in the target *of the right type* for the model to represent its target at all? For example, shall we correctly interpret the property "being an electron", but allow the electron of Rutherford's model to denote a hockey puck (implying that the model is not veridical, because a hockey puck is not an electron)? Or shall we require that the electron of the model be mapped to a real electron?

If we adopt the latter requirement, we encounter problems with theories that fail to refer. Presumably, we would like to say that phlogiston theory is *false*, not that it fails to represent anything (it can represent combustion phenomena for instance). But if we expect objects in a model to refer to objects of the right type in the target for the model to represent anything, then a model with non-existent types of objects such as phlogiston will fail to represent anything.

So maybe we should be liberal and accept that particular objects can be mapped in any way without restriction. We would only require that properties, relations and functions be interpreted in the right way. Then phlogiston theory would misrepresent, because there cannot be a mapping that renders its models veridical (because phlogiston does not exist), but it would still represent. But if we adopt this stance, we encounter another problem: theories will generally be false about everything that is outside of their domain of application. Because of our permissiveness, we fail to correctly delimit the domain of application of the theory, that is, the set of targets that it can legitimately represent, and about which we want the theory to be either true or false. We do not want to say that a theory of optics misrepresents the behaviour of fluids. We want to say that it does not represent the behaviour of fluids at all. This is not a problem for theories with a universal domain of application (the fundamental theories of physics?), but limiting potential truth to these theories only is problematic.

The solution to this dilemma for the semantic realist must be something along these lines: there should be a distinction between two kinds of properties. Some properties are required to correctly identify the domain of application of the theory, and to say whether a model really represents the target; in our examples, properties such as "water" or "combustion phenomena". These properties typically categorise phenomena of interest. Other properties are posited by the theory to explain these phenomena: in our examples, "H_2O" and "phlogiston". Assuming such a distinction, we should only expect objects of the model that instantiate the former kind of property to be mapped to real objects of the right type, for a model to represent its target. That is, models must be applied to the right types of objects, but only for *some* properties: the ones that identify the types of objects that the model is really "talking about".

But this solution, which, I think, is the right one, generates troubles for the semantic realist. Presumably, the distinction between these two kinds of properties is rooted in the fact that some superficial properties are more directly accessible to epistemic agents of our community than others, and that these are the properties involved in the specification of the domain of application of the theory. They concern targets that can be potentially denoted by epistemic agents. So this distinction is epistemically informed. Now the question is: how can we come up with a notion of theoretical truth that is not epistemically constrained if it rests on a distinction between the properties that are epistemically accessible and those that are not (or are, but less directly)?

This is the problem I will address in the next section.

5 Are Conditions of Application Epistemically Constrained?

To sum up what has been said so far, if we want to make sense of our notion of theoretical truth, we need to specify correctly the way the models of a theory are applied. This means adopting conceptual restrictions on the way models can be interpreted in terms of their targets. Presumably, these restrictions must take the form of an interpretation of part of the vocabulary used that fixes its extension in the world. And importantly, they should concern the correct identification of objects whose types are epistemically accessible, but not the hypothetical entities postulated by the theory, if we want to be able to say that a theory that does not refer is false, but still applicable.

Now here is a question: are these conditions of applicability of the theory supposed to be analysed in epistemic terms? Are these conditions to be understood in terms of possible measurements, possible observations, possible intentions or interventions of a member of our epistemic community? Or should they be expressed in non-epistemic terms, that is, only in terms of the model and the target?

Assuming that representation is indeed a relation between a vehicle, a target and a user, it seems sensible to interpret potential representational force in epistemic terms. After all, the point of representational force is to account for the possibility of misrepresentation, and by this, we mean that model *users* could be wrong. This account is achieved by distinguishing norms of application, by which the representation relation takes place (representational force) and conditions of accuracy. But sound norms must be applicable in principle, and for this, they must concern aspects that are epistemically accessible to the users, at least in principle. We would naturally expect that a scientist is capable of knowing that her model represents a target, even if she does not know whether it represents it accurately. This implies that we make a distinction between aspects that are more or less directly accessible to epistemic agents and aspects that are inaccessible or less directly accessible. Only the former should be mentioned in norms of applicability.

This is not to say that we should come up with correspondence rules or strict operational definitions that would tell whether a model applies to a target. It does not even mean that whether and how the model applies must be known a priori by users of the theory. They could defer this knowledge to experts, or the best way to operationalise linguistic terms in concrete applications could be discovered by experience, as a matter of coherence between the conceptual structure of the theory and practical constraints. But this means that whether a model applies must be robustly assertible in principle by members of our community.

Note that we are not expecting veridicality to be assertible in the same way, since we are concerned about semantic realism, not epistemic realism. Theoretical truth might well remain out of reach. But not the fact that a model is a legitimate (if perhaps bad) representation of its target to start with.

From these considerations, there is a great temptation to simply claim that the interpretation involved when applying a model to a target is really a mapping between parts of the model and potential observations, potential measurements on the target, etc., that is, between the model and aspects that are better described in epistemic terms. This is a sure way to ensure epistemic accessibility. But then the following question arises: can we make sense of a notion of theoretical truth that is not epistemically constrained, even assuming that conditions of application are epistemically constrained? And if not, how can we pretend to be semantic realists?

Now the fact that conditions of application depend on our epistemic position does not mean that they cannot be analysed in ways that are independent of our epistemic position. So there are other options for the semantic realist, which I will examine in due course. But let us first examine in more detail what is at stake if we "bite the bullet".

The problem for semantic realism is that the mapping we are talking about, the one that is involved when applying a model to a target, is what will determine whether the model is veridical or not: a model can only be veridical relative to a given mapping between model and target properties. Changing the mapping, or in our case, the epistemically constrained aspects of the target that the model represents, would alter the veridicality of the model, or the truth-conditions of its cognitive content: it would change *what we are talking about.* This entails that real entities are not the only truth-makers of the theory, but epistemically constrained aspects are involved as well.

This problem is particularly salient if the model is only interpreted in epistemically constrained terms, and if other theoretical properties are interpreted structurally. This is the case in particular for epistemic structural realism, in which case the problem reduces to a version of Newman's objection against stuctural realism: if the cognitive content of a model comprising H_2O molecules is only a structure, and if the terms "Hydrogen" and "Oxygen" are only structurally interpreted, then basically, the model is veridical if the epistemically accessible characteristics of water respect certain relations, which means that veridicality is exactly what an empiricist would call "empirical adequacy". This cannot be semantic realism.

But the problem is more general, and affects other positions as well. For example, an ontic structural realist can claim that the kind of relations she is talking about are "modal". However, even if the qualitative term "modal" were enough to make a relation count as "real" so as to avoid Newman's objection (but see Ruyant, 2019), the cognitive content of the theory cannot be understood *exclusively* in realistic terms, because these are modal relations between entities that are epistemically identified. And even someone assuming that models posit real entities, such as molecules or phlogiston, has to assume that the cognitive content of a theory, what makes the theory true, is really about how these real entities relate to the epistemic entities that enter into the mapping. The notion of truth involved is still epistemically constrained, even if not entirely: there is no way the theory can be interpreted without mentioning epistemically loaded terms.

Going back to our definition of semantic realism, recall that a theory is true if for any targets in the domain, there is at least one model and one licensed interpretation for which the model is veridical: $(\forall t)(\exists M \in T)(\exists a \in A(M,t))V(M,a,t)$. The problem is simply that if A is expressed in epistemic terms, then truth is epistemically constrained. Our theory is not only about reality, it is also about our observations, measurement results, information, intentions, or whatever epistemic entity is employed.

We could consider that since the conceptual norms that constrain applicability are at a theory level, the definition could be turned into $Truth(T,A) : \Phi(T,A)$, where A stands for an interpretation of the relevant vocabulary. This is an interesting move, but now the theory we are realist about is not T, but $T + A$. All this move does is make explicit the fact that our theory was never only about reality, but also about the epistemic entities that figure in A.

6 Externalism and Natural Kinds

As I mentioned earlier, there is a way out for the realist, which is to deny that applicability constraints should be analysed in epistemic terms, even if they are relative to our epistemic position.

The idea is the following: the fact that some characteristics of targets, the ones that allow us to identify them, are accessible to us, and that we choose to map the objects having these properties to objects having the right properties in our models, should be understood as a contingent fact about us and these properties from a realist perspective. But the objects and properties that are mapped are nevertheless real (as opposed to properties such as colours, which, arguably, are relative to us).

For example, the realist could maintain that water is identified by ostentation, by referring directly to the kind of substance that water is (following Kripke, 1980). Any theory of water is true if this kind is indeed as the theory describes it, for example, if it is composed of hydrogen and oxygen atoms (however these terms are interpreted). Assuming that water is H_2O, a model of the theory making use of the term “water” applies to a target only if this target is composed of H_2O molecules, but this does not require that the model describes the target as being composed of H_2O molecules; only that it employs the term “water”, which denotes this kind of substance. So the criterion of possible misrepresentation is fulfilled.

This is, I think, the only solution for the semantic realist. But this implies that the weakened versions of realism, such as structural realism, that attempt to dispense with reference to natural kinds (Ladyman and Ross, 2007; French, 2014), will simply not do: they are not semantic realist positions at all, because they are unable to express the cognitive content of theories in terms that are not epistemically constrained, otherwise they end up being trivially true. The "modal structure of the world" cannot float freely: it must be anchored to real properties, at least the ones that enter into conditions of application. The structural realist could argue that we successfully refer to "real patterns", but arguably, these patterns are identified in relation to us if they are identified at all, so we cannot dispense with epistemic notions (it's worth noting here that Ladyman and Ross claim that their version of structural realism is a modal empiricism).

Note that this problem, although similar to Newman's objection against structural realism, is distinct: the problem is not that a scientific model would fail to make any substantial claim about its target (because any structure is realized if the target has the right cardinality), but that scientific models would fail to have well-defined targets to start with, unless natural kinds are involved.

The idea that the mapping between our theories and the world should be expressed in realistic terms is rather puzzling in some respects. It assumes that we correctly "carve nature at its joints" when delimiting the domains of application of our theories. If we didn't, we would fail to represent anything without even knowing it, even if we constantly apply our theories. But as we have observed earlier, it is problematic to impose norms on users when they cannot be followed. So the semantic realist must explain how the fact that we refer to real entities rather than artefactual ones when applying our theories can be assertible in principle (arguably, real properties must not be to scarce for it to be possible). In this context, there seems to be no principled reason not to assume a natural kind interpretation for all theoretical vocabulary, but this means that there is no simple way to avoid the arguments that have been mounted against the idea that theories successfully refer, such as the pessimistic meta-induction (Laudan, 1981). Furthermore, there are examples of reference failure at the observational level, for example, in the identification of biological species, or in the case of nephrite and jadeite. Should we say that a model of jade does not represent anything, even though we can easily give the set of objects to which it applies?

The view also entails that we are in no position to make explicit the "real" conditions of application of our own theories, nor to express what it really means for our theories to be true, unless we already have a *true theory* with regards to its domain of application. But then truth is quasi-tautological. For example, the models of a theory of water should be mapped to substances of composition H_2O, and the theory is true if what it says about H_2O is true–this, we can say only because we know that water is H_2O. Or so we assume. This looks like putting the cart before the horse.

On the other hand, if we do not want to assume a priori that we have reached some kind of truth about the domain of application of our theories, we must say that the cognitive content of our own theories is not transparent, but then it seems that the norms of correct representation cannot really be followed, which is hard to swallow (this is strongly related to Dummett (1978)'s criticism of semantic realism). We don't really know what our theories are talking about. All we can do is gesture at it ("water is *this* substance"), but saying more would require using epistemic notions. So our theories might have "real" truth-conditions (or not if we fail to refer), but our actual grasp of their content is still epistemically constrained. This looks like a Pyrrhic victory for semantic realism: why not simply claim that our theories are about these epistemically constrained entities?

If this was not enough, remember some of the assumptions we have made along the way (section 3): that there are well-defined applicability and veridicality conditions for models outside of any context of use, and that these are well-defined even for "de-idealised" models that would necessitate infinite cognitive capacities to be manipulated.

In my view, all this constitute good reason to give up semantic realism and look for alternatives. The resulting positions might still qualify as realist if one is willing to apply the term in a liberal way (as in: internal realism), as a label of seriousness, as it were, but this is merely a terminological issue.

7 Pragmatist Alternatives

In conclusion, let me briefly explore some alternatives to semantic realism.

Many authors have defended broadly pragmatist views of science (see Winther, 2015, section 4). The main characteristic of these approaches is an emphasis on scientific practice and on the contextual use of theories and models, rather than on formal aspects and abstract relationships between theories and the world. In light of the discussion above, it might be possible to flesh out this general stance (or at least one version of it) as a semantic thesis that is distinct from semantic realism.

There are many more or less radical ways one can depart from semantic realism. One can deny that scientific theories are truth-bearers at all, and instead claim that they are either good or bad, just as tools can be (eliminative instrumentalism). Or one can claim that the truth-value of theories is detached from any notion of model veridicality, and that theories are generally true as a matter of convention, like grammatical rules (conventionalism). The idea that scientific models are "autonomous agents" (Morgan and Morrison, 1999), or that theories are tools for model construction (Suárez and Cartwright, 2007), developed in the pragmatist tradition, could support such views. Another possibility is to deny that theoretical truth goes beyond actual uses of the theory, in the spirit of verificationism, but this kind of position might be unattractive. In this section, I wish to explore pragmatist positions that share as much as possible with semantic realism, and in particular, the idea that theories can be true or false, and that this has to do with model veridicality, including when it comes to models that are not actually used, but *could* be used to represent. So I will leave out the other options in what follows.

I think the most direct way to implement the emphasis on context and use that characterises pragmatism is to take the notion of applicability introduced in section 3 to be contextual. This idea has intuitive appeal: whether a model is legitimate or not to represent a given target seems to depend on the context, for example, on the activity (predicting or explaining), on standards of accuracy and on objects and properties of interest. The latter aspect is particularly relevant in the case of idealisations, since we would expect that an idealised model that neglects some aspects of the target, for example, the friction of a plane, is not applicable in a context where scientists are interested in these aspects. The pragmatist approach promises a fine analysis of applicability that the semantic realist cannot afford. The idea can be formalised by allowing our function A to take the possible context of use as a parameter (where a possible context could perhaps be partially formalised in terms of salient objects and properties and degrees of accuracy, using coarse-grained properties such as intervals for quantities).

An immediate advantage of this approach is that we do not need to assume that idealised models have a de-idealised version, one that might be impossible to handle with finite cognitive capacities, since idealised models can be considered veridical relative to a context, that is, for relevant variables and standards on accuracy. Let us assume, following Bokulich (2016), that fictions and idealisations capture "patterns of counterfactual dependence" between relevant variables. This could be taken to be the defining characteristic of veridicality, once the relevant variables have been specified by the context of use. All we need to assume, to avoid trivialising truth, is that a true theory would also have a veridical model in contexts with higher standards of accuracy, and with other relevant properties.

Defining the range of possible contexts could be problematic. For example, how high possible standards of accuracy can be? A solution is to make theoretical truth relative to epistemic communities and to their technical abilities. This could be a way of understanding the position known as perspectival realism (for example, Massimi (2018) talks of "standards of performance-adequacy" being relative to epistemic communities).

This gives us the following definition for theoretical truth:

$$Truth(T) : (\forall C, t)(\forall M \in T)(\forall a \in A(M, C, t))V(M, a, t)$$

A theory is true if, whatever the target and context of use, its models that apply to this target in this context are veridical. This definition captures a pragmatic notion of truth in terms of ideal success for a theory.

Let us examine how the approach fares with respect to the problems mentioned in section 4. There must be constraints on which interpretations are licensed in a given context, so that theories are not trivially false, but these constraints must not be as strong as veridicality. These constraints can be expressed, as was suggested, in terms of a distinction between two types of objects and properties: the ones that matter for applicability (for which objects in the model must be mapped to objects of the right type) and the ones that only matter for veridicality. But now this distinction can be contextual: the objects and properties that matter for applicability are those that are salient in a given context, because they are the objects and properties the users of the theory are interested in accounting for. This guarantees their principled accessibility. And of course, the pragmatist need not shy away from interpreting model properties, in context, in epistemic terms, for example, in terms of stabilised measurement outcomes (the a-contextual "meaning" of theoretical terms being captured by a function from context to interpretation, as given by A) .

Pragmatism so conceived of does not require giving up truth-conditional semantics for theories, nor the idea that truth somehow depends on what the world is like, although it also depends on our epistemic perspective. Does it count as a "literal" interpretation of scientific discourse? The idea that literalness would be reserved to non-epistemic conceptions of truth is disputable. After all, as I hope to have shown in this paper, no definition of theoretical truth can claim to be particularly intuitive and natural in the semantic conception of theories. Scientists are presumably competent interpreters of their own theories, but at times you can hear physicists claiming that Newtonian mechanics is true "within its domain of validity", which is not clearly compatible with semantic realism. So if the idea is to make good sense of scientific discourse in general, semantic realism might not be the best option. Also note that the standard textbook formulation of quantum theory explicitly mentions measurement, an epistemically loaded notion, while so-called realist interpretations attempt to complete the standard formulation with additional structure (particles, collapses or alternative worlds (Maudlin, 1995)). It is far from clear who, in these matters, is really interpreting the theory "literally", and who is imposing kosher reinterpretations...

The main obstacle to adopting a pragmatist stance could be that it apparently clashes with a naturalistic world-view. But who ever claimed that epistemic or mind-dependent notions are unnatural? Perhaps pragmatism assumes that they are not reducible (that they cannot be reflexively represented, or not completely). But attempts to reduce them to, say, physical descriptions are plagued with difficulties. The only world-view that needs to be abandoned is the idea that there is a "view from nowhere", that is, that there can be representational content without any tie to an epistemic agent, not even an abstract one. Abandoning this idea is merely a mark of modesty.

References

Ainsworth P (2009) Newman's objection. British Journal for the Philosophy of Science 60(1):135–171

Bell J (2004) Against "Measurement'. In: Speakable and Unspeakable in Quantum Mechanics, Cambridge University Press, pp 213–231

Boesch B (2017) There Is a Special Problem of Scientific Representation. Philosophy of Science 84(5):970–981

Bokulich A (2016) Fiction As a Vehicle for Truth: Moving Beyond the Ontic Conception. The Monist 99(3):260–279, DOI 10.1093/monist/onw004

Bolinska A (2013) Epistemic representation, informativeness and the aim of faithful representation. Synthese 190(2):219–234, DOI 10.1007/s11229-012-0143-6, URL http://link.springer.com/10.1007/s11229-012-0143-6

Bueno O, French S (2011) How Theories Represent. British Journal for the Philosophy of Science 62(4):857–894

C A Da Costa N, French S (2003) Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning. Oxford University Press USA

Cartwright N (1983) How the Laws of Physics Lie, vol 34. Oxford University Press

Chakravartty A (2001) The semantic or model-theoretic view of theories and scientific realism. Synthese 127(3):325–345

Contessa G (2007) Scientific representation, interpretation, and surrogative reasoning. Philosophy of Science 74(1):48–68

Demopoulos W, Friedman M (1985) Bertrand Russell's the Analysis of Matter: Its Historical Context and Contemporary Interest. Philosophy of Science 52(4):621–639

Dummett M (1978) Truth and Other Enigmas, vol 31. Harvard University Press

Elgin M, Sober E (2002) Cartwright on Explanation and Idealization. Springer Netherlands, Dordrecht, pp 165–174, DOI 10.1007/978-94-017-1009-1_9

van Fraassen B (1980) The Scientific Image. Oxford University Press

van Fraassen B (2008) Scientific Representation: Paradoxes of Perspective, vol 70. Oxford University Press

van Fraassen B, Ransanz P (1985) On the Question of Identification of a Scientific Theory (A Reply to "Van Fraassen's Concept of Empirical Theory" by Pérez Ransanz). Critica 17(51):21 – 29

French S (2014) The Structure of the World: Metaphysics and Representation. Oxford Univeristy Press, Oxford, United Kingdom

French S, Ladyman J (1999) Reinflating the semantic approach. International Studies in the Philosophy of Science 13(2):103–121

French S, Saatsi J (2006) Realism about Structure: The Semantic View and Nonlinguistic Representations. Philosophy of Science 73(5):548–559

Frigg R (2006) Scientific representation and the semantic view of theories. Theoria 21(1):49–65

Giere R (1991) Explaining Science: A Cognitive Approach. Philosophical Review 100(4):653–656

Giere R (2010) Scientific perspectivism, paperback ed edn. Univ. of Chicago Press, Chicago, oCLC: 934432475

Halvorson H (2012) What Scientific Theories Could Not Be. Philosophy of Science 79(2):183–206

Kripke S (1980) Naming and Necessity. Harvard University Press

Ladyman J, Ross D (2007) Every Thing Must Go: Metaphysics Naturalized. Oxford University Press

Laudan L (1981) A Confutation of Convergent Realism. Philosophy of Science 48(1):19–49

Lutz S (2015) What Was the Syntax-Semantics Debate in the Philosophy of Science About? Philosophy and Phenomenological Research 91(2)

Massimi M (2018) Four Kinds of Perspectival Truth. Philosophy and Phenomenological Research 96(2):342–359, DOI 10.1111/phpr.12300, URL http://doi.wiley.com/10.1111/phpr.12300

Maudlin T (1995) Three Measurement Problems. Topoi 14(1):7–15

McMullin E (1985) Galilean idealization. Studies in History and Philosophy of Science Part A 16(3):247–273, DOI 10.1016/0039-3681(85)90003-2

Miller A (2003) The Significance of Semantic Realism. Synthese 136(2):191–217, DOI 10.1023/A:1024742007683, URL http://link.springer.com/10.1023/A:1024742007683

Morgan M, Morrison M (1999) Models as Mediators: Perspectives on Natural and Social Science. Cambridge University Press

Psillos S (1999) Scientific realism: how science tracks truth. Philosophical issues in science, Routledge, London ; New York

Putnam H (1980) Models and reality. Journal of Symbolic Logic 45(3):464–482

Ruyant Q (2019) Structural Realism or Modal Empiricism? British Journal for the Philosophy of Science

Saatsi J (2016) Models, Idealisations, and Realism. Springer International Publishing, Cham, pp 173–189, DOI 10.1007/978-3-319-28163-6_10

Shalkowski SA (1995) Semantic Realism. Review of Metaphysics 48(3):511 – 538

Suppe F (1989) The semantic conception of theories and scientific realism. University of Illinois Press, Chicago

Suppes P (1960) A comparison of the meaning and uses of models in mathematics and the empirical sciences. Synthese 12(2-3):287–301

Suárez M (2003) Scientific representation: Against similarity and isomorphism. International Studies in the Philosophy of Science 17(3):225–244

Suárez M (2004) An inferential conception of scientific representation. Philosophy of Science 71(5):767–779

Suárez M (2015a) Deflationary representation, inference, and practice. Studies in History and Philosophy of Science Part A 49:36–47, DOI 10.1016/j.shpsa.2014.11.001, URL https://linkinghub.elsevier.com/retrieve/pii/S0039368114001095

Suárez M (2015b) Representation in Science, vol 1. Oxford University Press, DOI 10.1093/oxfordhb/9780199368815.013.25, URL http://oxfordhandbooks.com/view/10.1093/oxfordhb/9780199368815.001.0001/oxfordhb-9780199368815-e-25

Suárez M, Cartwright N (2007) Theories: Tools versus models. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 39(1):62–81

Thomson-Jones M (2012) Modeling without Mathematics. Philosophy of Science 79(5):761–772

Winther RG (2015) The Structure of Scientific Theories. Stanford Encyclopedia of Philosophy