Scientific Representation and Theoretical Equivalence

James Nguyen*†

In this article I connect two debates in the philosophy of science: the questions of scientific representation and both model and theoretical equivalence. I argue that by paying attention to how a model is used to draw inferences about its target system, we can define a notion of theoretical equivalence that turns on whether models license the same claims about the same target systems. I briefly consider the implications of this for two questions that have recently been discussed in the context of the formal philosophy of science.

1. Introduction. Here is one question: in virtue of what do models represent their target systems? Or more specifically, in virtue of what do models represent their target systems in a way that allows us to reason about target systems via reasoning about their models? Here is another: under what conditions are two scientific theories (whatever they are) equivalent? Despite the heated debates surrounding these questions, investigations into how they relate (e.g., how insights from one can be used in the context of another) are rare, if not nonexistent. This is surprising. If scientific theories are, in some sense, composed of models that represent the world, then investigating how models from two distinct theories represent would seem to have implications on whether the theories are equivalent. I investigate these implications here.

*To contact the author, please write to: History and Philosophy of Science, University of Notre Dame, Notre Dame, IN 46556; e-mail: jnguyen4@nd.edu.

†I have benefited from numerous fruitful discussions with Roman Frigg and Bryan Roberts on the material contained in this article. I am also grateful to Marton Gömöri, Ashton Green, Xavi Lanao, Pablo Ruiz de Olano, and Jim Weatherall for very useful comments on previous drafts. Thanks also to audiences at the University of Notre Dame and the 2016 Philosophy of Science Association meeting for constructive feedback.

Philosophy of Science, 84 (December 2017) pp. 982–995. 0031-8248/2017/8405-0015$10.00
Copyright 2017 by the Philosophy of Science Association. All rights reserved.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

In section 2 I provide some context by briefly outlining each of the questions mentioned above. In section 3 I argue that what I take to be the most promising accounts of scientific representation provide a natural sense in which two models have the same 'representational content' and that such content is a function on facts about models combined with interpretational schema from these facts to claims to be imputed onto their target systems. I generalize this to the representational content of theories. In section 4 I discuss the impact of introducing these model fact to target claim schema into an account of model and theoretical equivalence. I point out two implications: first, how we should understand Weatherall's (forthcoming) recent attempt to diffuse the hole argument with reference to the representational capacities of Lorentzian manifolds and, second, how recent attempts to provide formal accounts of theoretical equivalence should be understood (Halvorson 2012, 2016; Barrett and Halvorson 2016; Weatherall 2016a, 2016b). The overall purpose of this article is to encourage a dialogue between philosophers working on formal aspects of scientific theories and philosophers working on more general questions concerning scientific representation.

2. Representation and Equivalence

2.1. Representation. Scientific models represent their target systems. This much is uncontroversial. How they do so is not. Two popular views are that scientific models are, in some sense, intended copies of their target systems. For the likes of Giere (2004) and Weisberg (2013), model users intend that their models are similar (in the relevant respects, to the appropriate degree) to their targets. If they are so similar, then the model is an accurate representation of its targets, at least with respect to the intended features and to the appropriate degree. For the likes of van Fraassen (2008) along with, perhaps,[1] French and his collaborators (e.g., Da Costa and French 2003; French 2003, 2014; Bueno and French 2011), models are mathematical structures, and model users intend them to be appropriately morphic to their target systems or data gathered from them. If such a morphism exists, then the models are accurate representations of their targets, at least with respect to the structural features of the target covered by the morphism in question.

Both approaches face serious challenges (Suárez 2003; Frigg 2006). And accounts of scientific representation that diverge from the copy-making accounts have emerged as attractive alternatives (Hughes 1997; Suárez 2004; Contessa 2007; Frigg 2010). According to Hughes's (1997, S328) "DDI" account, models denote their targets and allow model users to perform dem-

1. I say 'perhaps' because they express reservations regarding whether user intentions should be built directly into an account of scientific representation (French 2003; Bueno and French 2011).

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

onstrations on models, whose results are then interpreted in terms of their target systems. Suárez (2004, 773) submits that models represent their targets only if their representational force "points" toward their targets and they allow competent model users to draw inferences about their targets. Contessa (2007, 57–58) argues that models denote their targets and that parts of models are interpreted as (purported) parts of their targets. Frigg (2010, 126–32) agrees about denotation and adds that models come with keys that allow users to "translate" model facts into target claims.

Although there are important differences between these 'interpretational' accounts, for my current purposes what is important is that they all agree that the primary purpose of scientific models is to license inferences about their target systems, and the way they do this is a function on model facts (Hughes's demonstration, the premises of Suárez's inferences, the model part of Contessa's interpretation, and the arguments of Frigg's keys) combined with intentional acts of model users interpreting these facts in terms of their target systems (Hughes's and Contessa's interpretations, Suárez's inferential schema used, and Frigg's keys). I do not argue for them here; I want to explore their insights for questions that invoke the representational capacities of scientific models.

2.2. Equivalence. When are two scientific theories, T_1 and T_2, equivalent? This, obviously, depends on what scientific theories are. If they are sets of formulas, formalized in an appropriate language, then they are *logically equivalent* if and only if (iff) their semantic closures are identical. But this requires that they are formulated in the same signature and thus might be too strict a criterion. Alternatively, they are *definitionally equivalent* iff for each term in the signature of T_1, a sentence can be constructed defining it in terms of the signature of T_2, and mutatis mutandis for each term in the signature of T_2, and the result of adding the respective definitions to each theory provides two expanded theories that are logically equivalent (Glymour 1970; see also Barrett and Halvorson 2016).

Alternatively, if two theories are construed as collections (in a sense to be specified below) of models, then an alternative approach is required. Halvorson (2012) provides three options available to someone who wants to define theoretical equivalence in this context. Two theories T_1 and T_2, construed as sets of models, are theoretically equivalent iff (*a*) T_1 and T_2 have the same cardinality, (*b*) there is a pointwise isomorphism between T_1 and T_2 (i.e., a bijection f such that for each model $M \in T_1$, $f(M) = M'$, where $M' \in T_2$, and M and M' are isomorphic), or (*c*) $T_1 = T_2$. Only *b* is remotely plausible, but it requires specifying the sense in which M and M' can be isomorphic to each other (see Halvorson 2012; Glymour 2013; Lutz 2017, for useful discussions about this and whether it can be done without reference to a linguistic signature).

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Halvorson (2012, 2016) has also recently suggested an exciting third sense in which theories can be seen as theoretically equivalent.[2] This can be applied in cases in which theories are construed linguistically or as collections of models. The idea in both cases is to take theories to be categories consisting of objects (formulas, models) and arrows (deductive relationships, homomorphisms, or elementary embeddings). Two theories T_1 and T_2 thus construed are categorically equivalent iff there is a functor F from T_1 to T_2 such that F is full, faithful, and essentially surjective.

Although there are clear differences between these approaches, what is important to note is that all of them focus on the formal features of the theories (construed as sets of models or formulas or categories) in question.[3]

3. Representational Equivalence. The accounts of how models represent discussed in section 2.1 can be used to provide an account of equivalence that diverges (at least prima facie) from the accounts outlined in section 2.2. However, this divergence should not be taken to indicate an essential tension between those accounts and the one outlined below (which I take to supplement rather than supplant the formal investigation). My aim in this article is more modest: I only hope to show the benefit of keeping interpretive questions in mind when thinking about theoretical equivalence.

We begin by considering how two models can be equivalent with respect to their representational content. I suggest that this has two individually necessary and jointly sufficient conditions. First, M and M' have the same target system(s). Although this seems obvious, even this condition may be missed if one focuses solely on the formal properties of the models in question. As van Fraassen (2014, 278) points out, we should not rule out cases in which two models satisfy the same equation but one of them is targeted at gas diffusion and the other at temperature distribution through time. Or we can consider models that satisfy an inverse square law, but one targets gravitational interaction between two massive bodies, and the other electrostatic interaction between two charged points. Whether two models have the same target systems already depends on factors beyond the formal properties of the models in question.

2. This idea has been applied to specific instances of theoretical equivalence by Rosenstock, Barrett, and Weatherall (2015) and Weatherall (2016a, 2016b).

3. Rosenstock et al. (2015) and Weatherall (2016a, 2016b) do refer to 'empirical equivalence' in addition to the formal relations between theories construed as categories, but their emphasis is still on the latter. Moreover, the sort of equivalence I am considering here goes beyond equivalence of empirical content (Sklar [1982] provides a useful discussion of the relationship between empirical and theoretical equivalence). I am grateful to Jim Weatherall for drawing my attention to this.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

In addition to having the same target system(s) two models need to 'tell us the same things' about those systems in order to be equivalent. In the accounts of representation discussed previously, this gets cashed out as M and M' licensing the same conclusions about those systems, where the set of conclusions licensed by a model M depends on (1) facts about M and (2) the interpretational schema used by the model user (or community thereof). An interpretational schema can be thought of as a function that takes model facts to target claims.[4] This function does not have to include every fact about M in its domain, nor does it have to be a (proposed) morphism or specified similarity relation between M and its target system(s).

A crucial implication of this is that in order to define, and use, an interpretational schema between model facts and claims to be imputed onto a target system, we need to be able to access facts about the model. It is no good to have a schema that says '*if* the model is such and such, *then* it represents its target as such and such' without being able to determine the truth of the antecedent. Since if a 'model' is not used to infer things about its target, then it is not really a model. This is the lesson we learn from both the 'intended copy making' and the interpretational accounts of scientific representation discussed above. As I discuss below, this has important implications for what adherents to the semantic view of theories should take models 'to be'.

So in order to be equivalent with respect to their representational content, two models, via their interpretational schema, have to license the same conclusions about the same target systems:

> **Model equivalence:** two models M and M' are equivalent with respect to their representational content iff (i) M and M' have the same targets and (ii) they license the same claim about those targets via their interpretational schema.

Assuming that, in some sense, theories should be thought of as collections (possibly categories) of models, this can be generalized to account for theoretical equivalence well:

> **Theoretical equivalence:** Two theories T and T' are representationally equivalent iff for each claim about a target licensed by a model of T there is a model of T' that licenses the same conclusion about the same target

4. I use 'model facts' liberally. These can include things like 'object a is in the extension of the relation R_i', 'R_1 and R_2 are the only relations defined on the model', or 'there is a morphism between the model and some other model in the theory considered as a category'.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

under the same circumstances and (ii) mutatis mutandis for every claim about a target licensed by a model of T'.[5]

This seems so easy that one might object that these definitions cannot be informative. Of course whether two theories or models are equivalent depends on whether they 'tell us' the same things about the same target systems; working out when this is the case is precisely the point in investigating the notion of theoretical equivalence.

This objection misses the point. First, the condition that two models have the same target system(s) requires a detailed investigation into what establishes the representational relationship (e.g., denotation or 'representational force') between models and their targets. But this does not depend on the formal properties of models alone, and as a fact of the matter, some models target some systems, and others do not. Although the question of whether two models have the same targets seems like an obvious one, it places a substantial condition on the notions of equivalence defined above.

Second, whether two models license the same claims about a given target system depends, at least on the accounts outlined above, on the interpretational schema associated with the models. The notion of an interpretational schema is defined in the abstract: a function from model facts to target claims. Understanding how a specific model, or theory, works will require a detailed investigation into the relevant model facts and the details of the schema itself (how does it associate model facts with target claims?).

The point is that, if we are interested in notions of equivalence that are stated at such a level of generality as to apply to diverse scientific fields, then our conditions on equivalence best be stated at the appropriate level of generality. I think the conditions above are stated at the appropriate level and, moreover, are substantial enough that the question of whether they are met in any particular case is nonvacuous and requires detailed investigation into the scientific practice associated with the case.

The resulting accounts of equivalence go beyond formal investigations. They provide a standard of equivalence that is sensitive to the pragmatic features of using models to represent the world. I think this is how it should be. To co-opt van Fraassen's (2014) example, an account of model equivalence that focused purely on the formal aspects of models would be analogous to an account of aesthetic equivalence that focused purely on the material constitution of works of art. Of course, in some sense, what makes the Mona Lisa the Mona Lisa is determined by facts about how oil paint is distributed across a specific block of wood. But this does not suffice to estab-

5. This way of approaching theoretical equivalence has much in common with Coffey (2014). The current discussion builds on his by emphasizing how recent discussions of scientific representation lead to a similar conclusion.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

lish the representational content of the picture. Moreover, a copy of the picture, an object with the same oil-on-wood facts, would only be equivalent (in the sense of representing the same subject in the same way) to the Mona Lisa if the same interpretational schema was used to interpret it. If, for example, the copy was demonstrated in the context of a forgery class, then it could be interpreted in a very different manner delivering very different facts about the world.

Having said that, I do not take my proposed accounts of equivalence to be in competition with the formal accounts of theoretical equivalence. In one sense they might be taken to address different questions: one concerning the technicalities of formal equivalence, the other an interpretive question concerning representational equivalence. But I take it that the former question is not motivated by technical considerations alone. The underlying idea seems to be that once we have established an account of formal equivalence, then this can be put to work to help answer the interpretive question. This idea can be seen in the likes of North (2009) and Curiel (2014) who disagree about whether Lagrangian or Hamiltonian mechanics is most appropriate for representing classical systems. My claim here is simply that to fully answer the interpretive question, in addition to the question of formal equivalence, we should keep interpretational schema, and coincidence with respect to target systems, in mind.

4. Fruitful Dialogue. In this section I discuss how paying attention to the fact that models have to be interpreted when thinking about whether they are equivalent will prove fruitful in contexts in which their representational capacities are important.

4.1. Weatherall and the Hole Argument. The hole argument runs as follows. Take a Lorentzian manifold, a model from general relativity. This is a structure of the form $\langle \mathcal{M}, g_{ab} \rangle$, where $\mathcal{M}$ is a four-dimensional differentiable manifold and g_{ab} a Lorentz-signature metric tensor. We suppose that $\mathcal{M}$ has a Cauchy surface S and a time orientation. Take an open region U to the time-like future of S, and let $\psi : \mathcal{M} \to \mathcal{M}$ be a diffeomorphism that is the identity for $\mathcal{M} \backslash U$ but is not the identity for U. Using the push forward of ψ to define a new metric tensor, $\psi_* g_{ab}$, we have two isometric Lorentzian manifolds $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$. We can now ask about what these structures represent. The problem arises if we take the following conditions to hold: $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ both represent a space-time ST, but $\langle \mathcal{M}, g_{ab} \rangle$ represents ST accurately, and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ does not. This substantivalist position introduces indeterminism into general relativity because we can suppose that both models are accurate representations of the past (i.e., the area of space-time represented by the "past" of S, in the sense determined by the time orientation), and yet they represent the future in in-

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

compatible ways (they cannot both be accurate representations thereof). Someone who denies the substantivalist position takes it that if two isometric Lorentzian manifolds represent the same space-time, then they do so equally accurately.

In an attempt to diffuse this argument, Weatherall (forthcoming) argues that as long as we are careful in how we understand how mathematical structures like Lorentzian manifolds represent, we can resist the idea of two Lorentzian manifolds such that one accurately represents *ST* and the other does not. To do so he invokes the premise that "interpretations of our physical theories should be guided by the formalism of those theories . . . [and] . . . insofar as they are so guided, we need to be sure that we are using the formalism correctly, consistently, and according to our best understanding of the mathematics" (2). He further claims that "if a particular mathematical model may be used to represent a given physical situation, then any isomorphic model may be used to represent that situation equally well" (4).[6] Applying this observation to the Lorentzian manifolds defined above, we arrive at the following conclusion: "once one asserts that spacetime is represent[ed] by a Lorentzian manifold, one is committed to taking isometric spacetimes to have the capacity to represent the same physical situations, since isometry is the standard of isomorphism given in the mathematical theory of Lorentzian manifolds" (15).

One way of interpreting this claim in the terminology of this article is that if two isometric Lorentzian manifolds represent the same physical situations, then one accurately represents that situation iff the other does. Under this reading, the argument runs as follows:

1. If two models M and M' represent the same target system(s) T, then if M and M' are equivalent in the sense specified by the mathematical theory of M and M', then M and M' are equally accurate representations of T (premise)
2. $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ are equivalent (isometric) in the sense specified by the mathematical theory of Lorentzian manifolds (premise)
3. $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ both represent *ST* (suppose)
4. $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ are equally accurate representations of *ST* (from 1, 2, and 3)

6. To clarify terminology here, I interpret Weatherall as claiming that if a particular mathematical model is used to represent a given physical situation, then any isomorphic model that also represents that physical situation may do so equally accurately. I sharply distinguish between 'represent' and 'accurately represent', but I think this is what Weatherall has in mind.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

Premises 2 and 3 are uncontroversial. Premise 2 is a mathematical fact, and the idea that both $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ represent *ST* seems benign.[7]

We might, however, put pressure on premise 1. To do so it is worth unpacking what it means for two models to be equally accurate representations of *T*. According to the interpretational accounts of scientific representation, this amounts to them licensing the same conclusions about *T*, which depends on the interpretational schema used to generate claims about *T* from facts about *M* and M'. In this framework the assumption underpinning premise 1 is that since $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ are isometric they should be interpreted in the same way. That is, given that they are isometric, we should restrict the domain of the interpretation schemas to the facts that they share and map these facts to the same claims about *ST*.

I think this claim is true, but only in contexts in which all we care about are the isometric invariant facts about *ST*. But as Roberts (2014) notes, if this argument is supposed to block the hole argument, then it must be the case that the disagreement between the substantivalist and her interlocutor cannot be formulated: "In particular [this] prohibits isometric Lorentzian manifolds from [accurately] representing distinct physical situations for any purpose whatsoever, not even to express manifold substantivalism for the purposes of reductio. It is this presumption, that isometric Lorentzian manifolds represent the same situation equally well 'for all purposes,' which is supposed to prohibit the (reductio) hypothesis that two isometric spacetimes can ever [accurately] represent different things" (6).

We are now in a position to deny premise 1. The claim is that (supposing that *M* and M' represent the same target system(s) *T*) if *M* and M' are equivalent in the sense specified by the mathematical theory of *M* and M', then *M* and M' are equally accurate representations of *T* in all contexts and for all purposes. But this requires that since $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ are isometric, any way of interpreting them must deliver the same claims about *ST*. And this need not be true. Even the very same structure can be interpreted in one way in one context and another way in another.

A familiar example illustrates this. Suppose that you are in London attempting to navigate using the tube map. Someone familiar with the tube map will adopt an interpretational schema toward the map that includes the

7. To avoid confusion, sometimes the debate is stated in terms of whether "isometric Lorentzian manifolds represent the same physical situations" (Earman and Norton 1987, 520), but I take it that we are supposed to interpret this claim as 'if isometric Lorentzian manifolds represent the same physical situations, then they do so equally accurately'. Once an account of scientific representation is adopted in which models represent only if model users use them to do so, then the claim that any of the infinite structures isometric to $\langle \mathcal{M}, g_{ab} \rangle$, including those never considered by anyone, represents *ST* becomes implausible.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

map's topological facts in its domain but no metric facts (of the map). Someone unfamiliar with the tube map may adopt an interpretational schema that includes these metric facts as well. Under the former, the map is an accurate representation of London with respect to its topological properties and remains silent with respect to its metric ones. Under the latter, the map is an accurate representation of London with respect to its topological properties and inaccurate with respect to its metric ones. There is a clear sense in which the very same object—the tube map—can be associated with different interpretational schema, such that under one it is an inaccurate representation (with respect to metric facts about London) and under another it is not (in virtue of remaining silent about them). Requiring that equivalent, let alone identical, mathematical structures be associated with the same interpretational schema is too demanding, and thus premise 1 is false.

Weatherall (forthcoming, sec. 5) seems sensitive to this objection. He states that in order to makes sense of the claim that isometric Lorentzian manifolds differ with respect to their representational accuracy the substantivalist "needs to stipulate what the additional structure might be and why we should think it matters. And it is difficult to see how this could be done in a mathematically natural or philosophically satisfying way [footnote omitted]" (16–17). I take it that 'additional structure' here refers to facts about $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ that are not preserved by the isometry, facts that would have to appear in the domain of the substantivalist's interpretation schemas associated with them (such facts do exist, $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ are not the same mathematical object by construction). But under this reading, what Weatherall's argument establishes is as follows. Although in the context of doing relativistic physics $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ will be interpreted in the same way, in the context of the hole argument, the substantivalist needs to specify the facts in which $\langle \mathcal{M}, g_{ab} \rangle$ and $\langle \mathcal{M}, \psi_* g_{ab} \rangle$ come apart and how, via the claims they are associated with by their respective interpretational schemas, they allow for one to represent *ST* accurately while the other does not. When phrased like this, Weatherall's argument is recast as challenge for the substantivalist, rather than 'blocking' the hole argument per se.[8]

4.2. Formalism and the Pragmatics of Representation. I hope to have shown that thinking about interpretational schemas associated with models, rather than focusing solely on their formal properties, will allow us to properly appreciate arguments that turn on how they represent. I now briefly dis-

8. I am grateful to Jim Weatherall for confirming in personal communication that this is his current position. One might argue that this 'blocks' the hole argument as traditionally understood. I do not take a stand on this issue.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

cuss how the notion of theoretical equivalence I introduced in section 3 relates to those outlined in section 2.2.

My point in this subsection concerns what 'models' should be taken to be in the context of thinking about theoretical equivalence in terms of the formal properties of models associated with theories. Recall the claim that an object (mathematical structure) is a scientific model only if it is used as such, and this requires offering an interpretational schema from model facts to target claims. In order to generate such a schema, model users have to access facts about their models. In the case of physical models, like the Phillips-Newlyn machine (Morgan 2012, chap. 5), this is relatively straightforward. In the case of mathematical models, less so. Facts about mathematical structures have to be determined in some way; they are not self-evident. One (but maybe not the only) way to think about how this is done is in terms of the deductive closure of the axioms that are used to pick them out. In the first-order case, since the logic is sound and complete, two sets of axioms $\mathcal{A}$ and $\mathcal{B}$ have the same deductive closure iff they are satisfied by the same class of structures $\mathcal{S}$. This allows us to identify the 'theory' with $\mathcal{S}$ itself and ignore the axioms used to pick them out as alternative 'intrinsic characterizations' of the same theory (French and Saatsi 2006).

But this approach breaks down once we move to higher-order theories. Since (e.g.) second-order logic is not complete, there exist alternative axiomatizations of the same class of structures that have different deductive closure. Corcoran (1980) provides the following example of two alternative categorical (in the model-theoretic sense) axiomatizations of $\langle \mathbb{N}, s \rangle$.

Let $\mathcal{L}$ be a first-order language with identity including a constant '0' and a monadic function 's'. Augment $\mathcal{L}$ with a single monadic second-order variable P. Then let $\mathcal{T} = \{(P0 \wedge \forall x(Px \rightarrow Psx)) \rightarrow \forall\, yPy, \forall x\, \forall y(sx = sy \rightarrow x = y), \forall x(sx \neq 0)\}$. The set of sentences $\mathcal{T}$ is categorically satisfied by $\langle \mathbb{N}, s \rangle$, where the interpretation function takes '0' to 0 and 's' to the successor function.[9] It is a categorical characterization of the natural numbers. Let $\mathcal{T}' = \{(P0 \wedge \forall x(Px \rightarrow Psx)) \rightarrow \forall yPy, 0 = 0, \ldots, s^n0 = s^n0, 0 \neq s0, \ldots, s^n0 \neq s^m0\}$ for all n, and $m \in \mathbb{N}$ such that $n \neq m$. Corcoran (1980) shows that $\mathcal{T}$ is also categorically satisfied by $\langle \mathbb{N}, s \rangle$ (with the same interpretation function as before).

Although $\mathcal{T}$ is incomplete (it does not prove all sentences satisfied by $\langle \mathbb{N}, s \rangle$), we can still use it to derive all sorts of interesting sentences that are satisfied by that structure. The set of sentences $\mathcal{T}'$, however, cannot even prove basic arithmetical truths. For example, $\mathcal{T}' \nvdash \forall x(sx \neq 0)$ since there is

9. Notice that an 'interpretation function' is a very different thing from an 'interpretation schema'. The former is a function from syntax to structure; the latter a function from model facts to target claims.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

no finite derivation from the axioms in $\mathcal{T}'$ to this sentence.[10] But this means that these alternative axiomatizations provide access to different facts about the same mathematical structure. And since the inferences that $\langle \mathbb{N}, s \rangle$ licenses about a target system depend on the facts about $\langle \mathbb{N}, s \rangle$ to which we have access, this means that under different axiomatizations $\langle \mathbb{N}, s \rangle$ could license different inferences about its target system(s). Which means that $\langle \mathbb{N}, s \rangle$ under one axiomatization is not, in general, equivalent with respect to its representational content to $\langle \mathbb{N}, s \rangle$ under another. As Corcoran puts it: "[there is a] vast difference between characterizing an interpretation [structure] and axiomatising its set of truths" (1980, 204).

As a corollary, a theory consisting of the singleton $\{\langle \mathbb{N}, s \rangle\}$ should not be identified with the structure (or set containing the structure) itself, at least from a representational perspective, since under different axiomatizations the structure licenses different inferences.[11]

This generates a problem for the adherent of the semantic view of theories, and I take it to be a serious problem rather than "tilting at windmills" (van Fraassen 2014, 279). Suppose one identified a theory as a class of models. Of course, to present such a theory, one will have to use language. And, if you like, models can be associated with an interpretation function for the language used to present them. So isomorphisms between models in different theories can be defined straightforwardly in terms of the relations and functions that interpret the same syntactic symbols. This avoids the issues raised by Halvorson (2012). The question, however, is what attitude we should adopt to the formulas that are used to present these models.

Glymour writes: "to be at all plausible, the semantic view must distinguish between the content of a theory (its class of models) and the means of characterizing that class—the theory expressed in some language" (2013, 288). But then the content of a theory is identified with the class of models, and the theory, as expressed in a language used to present this class, is accorded the role of presenting theoretical content, rather than being included in it. I think that the above example shows that even this position is untenable in the context in which higher-order logics are used. If theoretical content somehow depends on representational capacity, and representational capacity somehow depends on the interpretational schema, and this some-

10. What governs $\vdash$ in second-order logic is contentious, but the result holds for any sound recursive system of deductions in the language under consideration. All this observation relies on is that $\vdash$ does not allow for infinite proofs. Since scientists do not have the time for them, this seems justified to me.

11. And the matter will not be improved if we consider theories as categories of structures since $\langle \mathbb{N}, s \rangle$ is the same under both axiomatizations. This category is straightforwardly categorically equivalent to itself. But, for the aforementioned reason, the inferences that this category licenses about target systems depends on the axiomatization used to pick it out.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

how depends on model facts we can access, then it remains that we cannot guarantee that the interpretational schema associated with $\langle \mathbb{N}, s \rangle$ under one axiomatization is also applicable to $\langle \mathbb{N}, s \rangle$ under another axiomatization. Sometimes we can present a class of models without being able to access all of the truths satisfied by those models. And these truths correspond to potential elements in the domain of interpretational schema. So it is not only that language must be used to present a class of models, but the language used to present that very class of models can constrain the interpretational schema available and therefore the representational content of the theory. Presentation matters, and it should be accorded a proper role even in a formal account of theoretical equivalence.

5. Conclusion. Mathematical structures, and models more generally, do not represent without some sort of intentional act of model users associating model facts with target claims. Once this aspect of scientific representation is made explicit, we are able to ask how it affects questions that rely on how models represent. I have argued that, at least under one way of reading it, Weatherall's recent attempt to dissolve the hole argument requires too strong a restriction on how we associate interpretational schema with mathematical structures. I have also suggested that formal approaches to theoretical equivalence could be augmented by investigating the interpretational schema associated with mathematical models and, moreover, that once we pay due heed to the limitations of how we present classes of models, the linguistic presentations themselves can be included in a theory's 'theoretical content'.

REFERENCES

Barrett, T. W., and H. Halvorson. 2016. "Glymour and Quine on Theoretical Equivalence." *Journal of Philosophical Logic* 45 (5): 467–83.

Bueno, O., and S. French. 2011. "How Theories Represent." *British Journal for the Philosophy of Science* 62 (4): 857–94.

Coffey, K. 2014. "Theoretical Equivalence as Interpretative Equivalence." *British Journal for the Philosophy of Science* 65 (4): 821–44.

Contessa, G. 2007. "Scientific Representation, Interpretation, and Surrogative Reasoning." *Philosophy of Science* 74 (1): 48–68.

Corcoran, J. 1980. "Categoricity." *History and Philosophy of Logic* 1 (1): 187–207.

Curiel, E. 2014. "Classical Mechanics Is Lagrangian: It Is Not Hamiltonian." *British Journal for the Philosophy of Science* 65 (2): 269–321.

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

Earman, J., and J. Norton. 1987. "What Price Spacetime Substantivalism? The Hole Story." *British Journal for the Philosophy of Science* 38 (4): 515–25.

French, S. 2003. "A Model-Theoretic Account of Representation; or, I Don't Know Much about Art . . . but I Know It Involves Isomorphism." *Philosophy of Science* 70:1472–83.

———. 2014. *The Structure of the World: Metaphysics and Representation*. Oxford: Oxford University Press.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).

French, S., and J. Saatsi. 2006. "Realism about Structure: The Semantic View and Nonlinguistic Representations." *Philosophy of Science* 73 (5): 548–59.

Frigg, R. 2006. "Scientific Representation and the Semantic View of Theories." *Theoria* 55 (1): 49–65.

———. 2010. "Fiction and Scientific Representation." In *Beyond Mimesis and Convention: Representation in Art and Science*, ed. R. Frigg and M. Hunter, 97–138. Berlin: Springer.

Giere, R. N. 2004. "How Models Are Used to Represent Reality." *Philosophy of Science* 71 (4): 742–52.

Glymour, C. 1970. "Theoretical Realism and Theoretical Equivalence." *PSA 1970: Proceedings of the 1970 Biennial Meeting of the Philosophy of Science Association*, 275–88. East Lansing, MI: Philosophy of Science Association.

———. 2013. "Theoretical Equivalence and the Semantic View of Theories." *Philosophy of Science* 80 (2): 286–97.

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

———. 2016. "Scientific Theories." In *The Oxford Handbook of Philosophy of Science*, ed. P. Humphreys, 585–608. Oxford: Oxford University Press.

Hughes, R. I. G. 1997. "Models and Representation." *Philosophy of Science* 64 (Proceedings): S325–S336.

Lutz, S. 2017. "What Was the Syntax-Semantics Debate in the Philosophy of Science About?" *Philosophy and Phenomenological Research* 95 (2): 319–52.

Morgan, M. 2012. *The World in the Model: How Economists Work and Think*. Cambridge: Cambridge University Press.

North, J. 2009. "The 'Structure' of Physics: A Case Study." *Journal of Philosophy* 106 (2): 57–88.

Roberts, B. W. 2014. "Disregarding the 'Hole Argument.'" Unpublished manuscript, PhilSci Archive. http://philsci-archive.pitt.edu/11687/.

Rosenstock, S., T. W. Barrett, and J. O. Weatherall. 2015. "On Einstein Algebras and Relativistic Spacetimes." *Studies in History and Philosophy of Modern Physics* B 52:309–16.

Sklar, L. 1982. "Saving the Noumena." *Philosophical Topics* 13 (1): 89–110.

Suárez, M. 2003. "Scientific Representation: Against Similarity and Isomorphism." *International Studies in the Philosophy of Science* 17 (3): 225–44.

———. 2004. "An Inferential Conception of Scientific Representation." *Philosophy of Science* 71 (Proceedings): 767–79.

van Fraassen, B. C. 2008. *Scientific Representation: Paradoxes of Perspective*. Oxford: Oxford University Press.

———. 2014. "One or Two Gentle Remarks about Hans Halvorson's Critique of the Semantic View." *Philosophy of Science* 81 (2): 276–83.

Weatherall, J. O. 2016a. "Are Newtonian Gravitation and Geometrized Newtonian Gravitation Theoretically Equivalent?" *Erkenntnis* 81 (5): 1073–91.

———. 2016b. "Understanding Gauge." *Philosophy of Science* 83 (5): 1039–49.

———. Forthcoming. "Regarding the 'Hole Argument.'" *British Journal for the Philosophy of Science*.

Weisberg, M. 2013. *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

This content downloaded from 128.041.035.055 on April 24, 2019 04:20:45 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).