Scientific Representation, Interpretation, and Surrogative Reasoning*

Gabriele Contessa†‡

In this paper, I develop Mauricio Suárez's distinction between denotation, epistemic representation, and faithful epistemic representation. I then outline an interpretational account of epistemic representation, according to which a vehicle represents a target for a certain user if and only if the user adopts an interpretation of the vehicle in terms of the target, which would allow them to perform valid (but not necessarily sound) surrogative inferences from the model to the system. The main difference between the interpretational conception I defend here and Suárez's inferential conception is that the interpretational account is a substantial account—interpretation is not just a "symptom" of representation; it is what makes something an epistemic representation of a something else.

1. The Inferential Conception of Scientific Representation. Most philosophers of science today agree that models play a central role in science and that one of their main functions is to represent aspects of the world (see Hughes 1997; French and Ladyman 1999; Giere 1999; Teller 2001; Suárez 2002; Bailer-Jones 2003; French 2003). R. I. G. Hughes, for one, claims, "The characteristic—perhaps the only characteristic—that all theoretical models have in common is that they provide representations of parts of the world" (Hughes 1997, S325). The Rutherford model of the atom, for example, represents the atom, and the ideal pendulum model can be used to represent the tire swing hanging from a tree in the garden. However, there is still much disagreement as to what we mean when we say that a model represents a system.

In this article, I follow Mauricio Suárez in distinguishing different notions of representation. In particular, Suárez distinguishes between rep-

*Received August 2005; revised January 2007.

†To contact the author, please write to: Department of Philosophy, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada; e-mail: gabriele_contessa@carleton.ca.

‡I would like to thank Nancy Cartwright and Mauricio Suárez for their helpful comments on previous versions of this article.

Philosophy of Science, 74 (January 2007) pp. 48–68. 0031-8248/2007/7401-0004$10.00
Copyright 2007 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/519478 Published online by Cambridge University Press

resentation and "accurate, true and complete representation" (Suárez 2004, 767). According to Suárez, the primary aim of a *substantial account* of scientific representation—that is, an account which specifies necessary and sufficient conditions for scientific representation—is to answer the question: "In virtue of what does a certain model represent a certain system?" and not the question: "In virtue of what does a certain model represent a certain system accurately or truthfully?" (Suárez 2004, 767–768).

According to the inferential conception of scientific representation that Suárez proposes, "*A* represents *B* only if (i) the representational force of *A* points to *B* and (ii) *A* allows competent and informed agents to draw inferences regarding *B*" (Suárez 2004, 773). According to Suárez's account, there are two conditions that a model has to satisfy in order to represent the system: the first condition, which I shall call *denotation*, is that the model is used by someone to represent the system; the second, which I shall call *surrogative reasoning*, is that the model allows its user(s) to perform specific inferences from the model to the system.

According to the inferential conception, however, denotation and surrogative reasoning, though necessary, are not jointly sufficient conditions for a model to represent a system. Therefore, the inferential conception is not a substantial account of representation. Suárez claims that "representation is not the kind of notion that requires, or admits, [universal necessary and sufficient] conditions" (Suárez 2004, 771). Thus, according to Suárez, an account of scientific representation cannot (and should not) provide us with a set of jointly sufficient conditions for scientific representation. As Suárez seems to acknowledge, though, his reasons for claiming this are ambiguous (see Suárez 2004, n. 4).

On the one hand, Suárez claims that one should not expect further conditions because there are "no deeper features to scientific representation other than its surface features" (i.e., denotation and surrogative reasoning) (Suárez 2004, 769). It is in this vein, I suspect, that Suárez claims that the inferential conception is a *deflationary* conception of scientific representation (Suárez 2004, 770–771).

On the other hand, Suárez thinks that additional conditions may need to hold in order for a model to represent a system, but that these additional conditions are different in different instances of scientific representation. For example, Suárez claims that: "In every specific context of inquiry, given a putative target and source, some stronger condition will typically be met; but which one specifically will vary from case to case. In some cases it will be isomorphism, in other cases it will be similarity, etc." (Suárez 2004, 776).[1] According to this interpretation, Suárez would be

1. Similarly, see also Suárez (2003, 768 and 776).

https://doi.org/10.1086/519478 Published online by Cambridge University Press

claiming that no *general* account of representation can specify necessary and sufficient conditions because, beside denotation and surrogative reasoning, there are no other general conditions that are sufficient for scientific representation. If this interpretation is correct, Suárez's inferential conception is a *minimalist* conception of representation rather than deflationary one.[2]

Suárez "propose[s] that we adopt from the start a deflationary or minimalist attitude . . . towards the concept of scientific representation" (Suárez 2004, 770). However, it is not clear why we should adopt a deflationary attitude *from the start*. The claim that representation does not admit universal necessary and sufficient conditions seems only indirectly supported by Suárez's arguments from the lack of success of two specific versions of the similarity and structural conceptions of representation, [iso] and [sim], to the correctness of a deflationary approach (see Suárez 2002 and 2003). However, these arguments are far from establishing that no adequate substantial account of scientific representation is viable.

First, insofar as [iso] and [sim] can be interpreted as either accounts of scientific representation or accounts of faithful scientific representation, they would seem to be accounts of faithful scientific representation. Even if they actually failed to provide us with an adequate substantial account of *faithful* scientific representation, this would hardly prove that it is not possible to formulate an adequate, substantial account of scientific representation *simpliciter*. Second, even if [iso] and [sim] were actually accounts of scientific representation *simpliciter*, these two accounts are far from being the only two possible substantial accounts of scientific representation.[3] Indeed, they are a far cry from exhausting the set of all the possibilities.

In this article, I will substantiate this claim by proposing and defending an account of scientific models as epistemic representations of systems in the world, which, I argue, is both adequate and substantial—that is, one that answers the question "*In virtue of what* does a certain model represent a certain system?" If I am right, this undermines Suárez's reasons for settling for an nonsubstantial account of scientific representation, for I take it that, if it is possible to provide both necessary and sufficient conditions for scientific representation, we should do so and not settle for an account that only provides a set of individually necessary but not jointly sufficient conditions.

2. This is the interpretation Suárez seems to favor in a recent joint paper with Albert Solé (Suárez and Solé 2006).

3. Most of Suárez's arguments against [sim], for example, do not seem to apply to the version of the similarity conception of scientific representation defended by Ronald Giere (2004) and Paul Teller (2001).

https://doi.org/10.1086/519478 Published online by Cambridge University Press

The importance of Suárez's work on scientific representation can be scarcely overestimated. In particular, Suárez has to be praised for making explicit the close connection between scientific representation and surrogative reasoning and for insisting on the importance of distinguishing between what I will call epistemic representation and faithful epistemic representation. However, Suárez's pessimism about finding deeper features of scientific representation leaves one with the mistaken impression that there is something mysterious about our ability to use models to perform pieces of surrogative reasoning about their target systems.[4] If the account I develop and defend here is correct in its main lines, the user's ability to perform surrogative inferences from the model to the system can be explained by the fact that the user interprets the model in terms of the system. Interpretation is what grounds both scientific representation and surrogative reasoning.

2. Surrogative Reasoning: Validity and Soundness. "Surrogative reasoning" is the expression introduced by Chris Swoyer (1991) to designate those cases in which someone uses one object, the *vehicle* of representation, to learn about some other object, the *target* of representation. A good example of a piece of surrogative reasoning is the case in which someone uses a map of the London Underground to find out how to get from one station on the London Underground network to another. The map and the network are clearly two distinct objects. One is a piece of glossy paper on which colored lines and names are printed; the other is an intricate system of, among other things, trains, tunnels, rails, and platforms. By examining the map, however, one can learn a great deal about the network. For example, one can find which trains to catch in order to reach one of the stations on the network from any other station on the network. We can thus say that the map allows its users to carry out a piece of surrogative reasoning about the network or, less awkwardly, that the users of the map can perform surrogative inferences from the map to the network.

Here, it is useful to sketch the distinction between valid and sound surrogative inferences (even if the resources needed to define the notion of valid surrogative inference will only become available later). A surrogative inference is *sound* if it is valid and its conclusion is true of the target. However, a surrogative inference can be valid even if it is not sound (i.e., an inference is valid irrespectively of the truth of its conclusion).

4. The sense of mystery about surrogative reasoning also haunts the somewhat similar account of scientific representation held by Daniela Bailer-Jones. According to her account, models "entail" propositions in some not-better-specified nonlogical sense (Bailer-Jones 2003).

https://doi.org/10.1086/519478 Published online by Cambridge University Press

3. Denotation, Epistemic Representation, and Faithful Epistemic Representation. One of the main problems with the notion of representation is that by "representation" people often mean different things. For the present purposes, it is important to distinguish three senses of "representation." In a first sense, both the logo of the London Underground and a map of the London Underground represent the London Underground network. In the terminology I adopt here, we can say that they both *denote* the network. Denotation may well be a matter of convention. Here, I will assume that, in principle, anything can denote anything else if a group of users implicitly or explicitly agree that it does so.

However, the map of the London Underground does more than just denoting the London Underground network; it represents the London Underground in a second, stronger sense: it is an *epistemic representation* of the network. It is in virtue of the fact that the map of the London Underground represents the London Underground in this stronger sense that a user can perform (valid though not necessarily sound) surrogative inferences from the map to the network. The same does not apply to the London Underground logo. If one has to figure out how to go from Holborn to Finsbury Park by tube, one can use the map, but there is no obvious way to use the logo. An examination of the logo does not allow us to infer much about the London Underground network. In the terminology I shall use here, the map is an *epistemic representation* of the network, while the logo is not.[5]

On the conception of epistemic representation that I will defend here, the fact that some user performs a surrogative inference from a certain object, the vehicle, to another, the target, is merely a "symptom" of the fact that, for that user, that vehicle is an epistemic representation of that target, a symptom that allows us to distinguish cases of epistemic representation (such as the London Underground map) from cases of mere denotation (such as the London Underground logo). However, there may be "asymptomatic" cases of epistemic representation. That is, a user does not necessarily need to perform any *actual* piece of surrogative reasoning in order for the vehicle to be an epistemic representation of the target. For example, even if someone has never performed and will never perform any actual inference from the London Underground map to the London Underground network, the map may still be an epistemic representation of the network for her provided that she would have been able to perform one of those inferences if only the occasion had arisen.

We can now formulate the following characterization of the notion of epistemic representation. A vehicle is an *epistemic representation* of a cer-

5. As far as I can see, this distinction coincides with the distinction drawn by Suárez between representation and cognitive representation (Suárez 2004, 772).

https://doi.org/10.1086/519478 Published online by Cambridge University Press

tain target for a certain user if and only if the user is able to perform valid (though not necessarily sound) surrogative inferences from the vehicle to the target. I will call this necessary and sufficient condition for epistemic representation *valid surrogative reasoning*. This definition thus says that valid surrogative reasoning is a necessary and sufficient condition for epistemic representation.

Two remarks are in order here. First, according to the characterization of epistemic representation above, a vehicle is not an epistemic representation of a certain target in and of itself—it is an epistemic representation *for someone*. Epistemic representation is not a dyadic relation between a vehicle and a target but a triadic relation between a vehicle, a target, and a (set of) user(s).[6] For the sake of simplicity, I will often omit to mention the users of an epistemic representation unless it is required by the context. However, this does not mean that a vehicle can be an epistemic representation of a target for no one in particular or in its own right—a vehicle is an epistemic representation of a certain target only if there are some users for whom it is an epistemic representation of that target.[7]

Second, the notion of epistemic representation is primarily a technical notion. As many technical notions, however, the notion of epistemic representation is meant to capture what I take to be one of the senses of the ordinary notion of representation. According to this definition, numerous prototypical cases of what we would ordinarily consider representations

6. That epistemic representation is not an intrinsic relation between a vehicle and a target but a triadic relation that involves a vehicle, a target, and a set of users seems to be one of the few issues on which most contributors to the literature on scientific representation agree (see, e.g., Frigg 2002; Suárez 2002 and 2003; Giere 2004). Suárez (2002), however, does not seem to think that this is the case. He thinks that the supporters of the similarity and structural conception of epistemic representation are trying to "naturalize" epistemic representation in the sense that they are trying to reduce representation to a dyadic relation between the vehicle and the target. Whereas some early work by, say, Ronald Giere and Steven French may at times give this impression, I do not think that these are Giere's or French's considered views. Giere has recently dispelled any doubt by declaring: "The focus on language as an object in itself carries with it the assumption that our focus should be on *representation*, understood as a two-place relationship between linguistic entities and the world. Shifting the focus to scientific practice suggests that we should begin with the activity of *representing,* which, if thought of as a relationship at all, should have several more places. One place, of course, goes to the agents, the scientists who do the representing" (Giere 2004, 743).

7. This is particularly important when the epistemic representation has a large set of users (such as the London Underground map). In those cases, we usually tend to disregard the fact that the vehicle is an (epistemic) representation *for those users* not *in its own right*. The fact that a vehicle is an epistemic representation for many people or even for everyone, however, does not imply that it is an epistemic representation in and of itself.

https://doi.org/10.1086/519478 Published online by Cambridge University Press

turn out to be epistemic representations for us. Portraits, photographs, maps, graphs, and a large number of other representational devices usually allow their users to perform (valid) inferences to their targets and, as such, according to the characterization above, they would be considered epistemic representations of their targets (for us). For example, according to that characterization, if we are able to perform (valid) inferences from a portrait to its subject (as we usually seem able to do), then the portrait is an epistemic representation of its subject (for us).

However, there is also a sense in which the notion of epistemic representation seems to be broader than the ordinary notion of representation. This, I think, is due to an ambiguity of the ordinary notion of representation. "Represent" is sometimes used as a success verb and sometimes not. This is probably why we usually tend to conflate two distinct facts—the fact that a certain vehicle is an (epistemic) representation of a certain target and the fact that it can be a more or less faithful (epistemic) representation of that target.

Consider, for example, an old 1930s map of the London Underground and a new map of the London Underground. Both represent the London Underground network in the sense that one can perform valid surrogative inferences from either map to the network and both represent the same aspects of the network. But they offer conflicting representations of some of these aspects. For example, from the old map, one would infer that there is no direct train connection between Euston and Oxford Circus, while, from the new map, one would infer that Victoria Line trains operate between these two stations.

Whereas all the valid inferences from the new map to the network are sound, some of the inferences from the old map to the network that are valid according to its standard interpretation will no longer be sound today because, in the meantime, the network has significantly changed. In this sense, only the new map of the London Underground *faithfully represents* today's network, while the old map's representation of it is only *partially faithful*, as it *misrepresents* some aspects of the network. (Obviously, the reverse is true of the London Underground network of the 1930s, which is faithfully represented by the old map but not by the new map.) In general, a vehicle is a *completely faithful representation* of a target if and only if the vehicle is an epistemic representation of the target and all of the valid inferences from the vehicle to the target are sound. It is a *partially faithful representation* of a target if and only if the vehicle is an epistemic representation of the target and some of the valid inferences from the vehicle to the target are sound. It is a *completely unfaithful epistemic representation* of a target if and only if the vehicle is an epistemic representation of the target and none of the valid inferences from the vehicle to the target are sound. A vehicle *misrepresents* (some aspects of)

https://doi.org/10.1086/519478 Published online by Cambridge University Press

a target if the vehicle is an epistemic representation of the target and some of the valid inferences from the vehicle to the target are not sound.[8] The distinction between epistemic representation and faithful epistemic representation is analogous to Suárez's distinction between scientific representation and "accurate, true, or complete" representation.

Unlike epistemic representation, faithful epistemic representation is a matter of degree. A representation can be more or less faithful to its target. The same vehicle can be a faithful representation of some aspects of the target and misrepresent other aspects. This seems to be the case with the old London Underground map. The map misrepresents today's system in the sense that, from it, it is possible to draw many false conclusions about it. However, from it, it is also possible to draw many true conclusions about aspects of today's network. Unlike the occasional tourist, a knowledgeable user may be able to use the old map successfully for many purposes even if the map, to a large extent, misrepresents the system. The knowledgeable user, for example, knows that if, on the old map, two stations are not connected by any line, she should not conclude that in today's network there is no direct train service between the two stations for some of the subway lines which operate in today's network did not exist in the 1930s.

The notion of completely faithful representation seems to coincide with what Suárez calls a true representation. It is important to note the probably obvious fact that, in order to be a completely faithful representation of its target, a representation does not need to be a perfect replica of its target. The new London Underground map is an example of a completely faithful representation (or, at least, I assume it is), but it is not a perfect replica of the London Underground network. There are innumerably many aspects of the London Underground network that are beyond the representational scope of the map (such as the internal structure of stations, the spatial relations among them, and so on).[9]

8. Let me note that only by drawing the distinction between epistemic representation and faithful epistemic representation we can explain why something can both represent and *mis*represent something else. According to the picture sketched here, a vehicle must represent (i.e., be an epistemic representation of) a target in order to *mis*represent it.

9. A *completely faithful* representation is thus to be distinguished from what Suárez calls a *complete* representation. Indeed, I doubt that any representation is complete in Suárez's sense. According to Suárez, "[A representation] is complete if it is true and fully informative, licensing inferences to all truths about the target" (2004, 776). As far as I can see, every real representation falls far short of being complete in Suarez's sense. All actual representations represent only some aspect or other of their targets. Even the map as large as the territory mentioned in Lewis Carrol's *Sylvie and Bruno* seems to fall short of being a complete representation in Suarez's sense as it cannot allow its user to infer *all* truths about the territory.

https://doi.org/10.1086/519478 Published online by Cambridge University Press

I suspect that the fact that "representation" is used to refer to what I call denotation, epistemic representation, and faithful epistemic representation is the source of many of the problems associated with the notion of "representation." Ordinary language does not distinguish between these different senses of "representation," and their conflation seems to be the source of many misunderstandings among philosophers who are interested in representation. I think that many misunderstandings could be easily avoided by carefully distinguishing between different senses of representation as Suárez proposes we do.

4. Scientific Models as Epistemic Representations. How do the notions I have introduced so far relate to the problem of how scientific models represent? Scientific models, I claim, are epistemic representations of aspects of certain systems in the world. A certain model represents a certain target primarily in the sense that a user can use the model to perform surrogative inferences from the model to the system.

The Rutherford model of the atom, for example, was proposed by Ernst Rutherford (1911) in order to account for the phenomenon now known as Rutherford scattering. In a series of experiments in 1909, Hans Geiger, one of Rutherford's collaborators, and Ernest Marsden, one of Geiger's students, found that, in passing through a foil of gold 0.00004 cm thick, one in 20,000 alpha particles was scattered at an average angle of 90° (Geiger and Marsden 1909). The phenomenon could not be accounted for by what, at the time, was the main model of the atom—the Thomson model of the atom, also informally known as the plum pudding model. In the plum pudding model, the negatively charged electrons are embedded in a sphere of uniform positive charge that takes up the whole volume of the atom, like raisins in a plum pudding. The positive charge and mass are uniformly distributed over the volume of the atom. If the golden foil in Geiger and Marsden's experiment were made up of atoms like the ones in the Thomson model, even if all of the approximately 400 atoms in the foil fortuitously happened to deflect an alpha particle in the same direction, the particle would still be scattered at a very small angle. From the Thomson model of the atom or, more precisely, from a model of Geiger and Marsden's experiment in which the atoms in the golden foil are represented as in the Thomson model of the atom, we can therefore infer that Rutherford scattering would never occur. However, since Geiger and Masden's 1909 experiments show that the phenomenon actually occurs, the Thomson model misrepresents (that aspect of) the atom, for it leads to a false conclusion about it.

In the Rutherford model, on the other hand, all the positive charge of the atom and almost all of its mass are concentrated in the nucleus, whose radius is one-hundredth that of the atom, and the rest of the volume of

https://doi.org/10.1086/519478 Published online by Cambridge University Press

the atom is empty except for the orbiting electrons. Since the total deflection of a positively charged particle by a sphere of positive charge increases as the inverse of the radius of the sphere, the encounter with one single nucleus can deflect an alpha particle at an angle of 90°. However, since most of the volume of the atom is empty except for the electrons, and the mass of electrons is too small to scatter high-momentum alpha particles, most alpha particles will not be deflected at large angles. Unlike from the Thomson model of the atom, thus, we can soundly infer that Rutherford scattering will occur from the Rutherford model.

When we say that the Thomson and Rutherford models of the atom represent the gold atom, we are not merely saying that they denote the gold atom, like "Au" on the periodic table does. Rather, we are saying that they are epistemic representations of the atom—in the sense that both can be used by their users to perform inferences about certain aspects of the atom.

5. Epistemic Representation and Analytic Interpretations. I will now put forward and develop a substantial conception of epistemic representation which I call the *interpretational conception of epistemic representation*. According to the interpretational conception, a vehicle is an epistemic representation of a certain target (for a certain user) if and only if the user adopts an interpretation of the vehicle in terms of the target. I will call this necessary and sufficient condition *interpretation*. But what is an interpretation?

According to a general, though somewhat loose, characterization of the notion of interpretation, a user interprets a vehicle in terms of a target if she takes facts about the vehicle to stand for (putative) facts about the target. One specific way to interpret the vehicle in terms of the target (though possibly not the only way) is to adopt what I will call an *analytic interpretation* of the vehicle in terms of the target. An analytic interpretation of a vehicle in terms of the target identifies a (nonempty) set of relevant objects in the vehicle ($\Omega^V = \{o_1^V, \ldots o_n^V\}$) and a (nonempty) set of relevant objects in the target ($\Omega^T = \{o_1^T, \ldots o_n^T\}$), a (possibly empty) set of relevant properties of and relations among objects in the vehicle ($P^V = \{{}^nR_1^V, \ldots, {}^nR_m^V\}$, where ${}^nR$ denotes an n-ary relation and properties are construed as 1-ary relations) and a set of relevant properties and relations among objects in the target ($P^T = \{{}^nR_1^T, \ldots, {}^nR_m^T\}$), and a set of relevant functions from $(\Omega^V)^n$—that is, the Cartesian product of Ω^V by itself n times—to Ω^V ($\Phi^v = \{{}^nF_1^V, \ldots, {}^nF_m^V\}$, where ${}^nF$ denotes an n-ary function) and a set of relevant functions from $(\Omega^T)^n$ to Ω^T ($\Phi^T = \{{}^nF_1^T, \ldots, {}^nF_m^T\}$).

A user adopts an *analytic interpretation* of a vehicle in terms of a target if and only if:

https://doi.org/10.1086/519478 Published online by Cambridge University Press

1. The user takes the vehicle to denote the target,
2. The user takes every object in Ω^V to denote one and only one object in Ω^T and every object in Ω^T to be denoted by one and only one object in Ω^V,
3. The user takes every n-ary relation in P^V to denote one and only one relevant n-ary in P^T and every n-ary relation in P^T to be denoted by one and only one n-ary relation in P^V,
4. They take every n-ary function in Φ^V to denote one and only one n-ary function in Φ^T and every n-ary function in Φ^T to be denoted by one and only one n-ary function in Φ^V.

Most interpretations of vehicles in terms of targets that we ordinarily adopt seem to be analytic. The standard interpretation of the London Underground map in terms of the London Underground network, for example, is an analytic interpretation. First, we take the map to denote the London Underground network—that is, we take the map to be a map of the London Underground network and not, say, the New York subway network. Second, we take small black circles and small colored tabs with a name printed on the side to denote stations on the network with that name. Third, we take some of the properties of and relations among circles and tabs on the map to stand for properties of and relations among stations on the network. For example, we take the relation *being connected by a light blue line* on the map to stand for the relation *being connected by Victoria Line trains* on the network.

For the sake of simplicity, here I will focus exclusively on analytic interpretations. So, when saying that a user adopts an interpretation of the target, I will always assume that the interpretation in question is an analytic one. However, I do not mean to imply that all interpretations of vehicles in terms of the target are necessarily analytic. Epistemic representations whose standard interpretations are not analytic are at least conceivable. However, in practice, epistemic representations whose standard interpretations are nonanalytic seem to be the exception rather than the rule. In the overwhelming majority of prototypical cases of epistemic representation (which include maps, diagrams, drawings, photographs, and, of course, models), it seems possible to reconstruct the standard interpretation of the vehicle in terms of the target as an analytic one.[10] If this is true, then restricting our attention to analytic interpretations will considerably simplify the discussion without any comparable loss of generality.

10. I talk of reconstruction because users are often unable to spell out how they interpret the vehicle in terms of the target and sometimes are not even aware that they do interpret the vehicle in terms of the target.

https://doi.org/10.1086/519478 Published online by Cambridge University Press

6. Interpretation and Surrogative Reasoning. According to the interpretive conception of representation I defend here, users adopt an interpretation of the Rutherford model in terms of an atom if and only if: (1) they take the model as a whole to stand for the atom in question (i.e., they take the model to be *a model of* the atom), (2) they take some of the components of the model to stand for some of the components of the system (e.g., they take the nucleus of the atom in the model to stand for the nucleus of the atom in the system, the electron of the atom in the model to stand for one of the electrons in the system), and (3) they take some of the properties of and relations among the objects in the model to stand for the properties of and relations among the corresponding objects in the system (if those objects stand for anything in the atom). So, for example, the fact that the nucleus in the model is positively charged stands for the fact that the object denoted by the nucleus, the nucleus of the atom in the system, is positively charged.

A few remarks are in order here. First, a user does not need to believe that every object in the model denotes some object in the system in order to interpret the model in terms of the system. For example, in the Aristotelian model of the cosmos, the universe is represented as a system of concentric crystal spheres. The Earth lies at the center of the sublunary region, which is the innermost sphere. Outside the sublunary region are the heavens: eight tightly fit spherical shells. The outermost spherical shell, the sphere of the fixed stars, hosts the stars. Each of the other spherical shell hosts one of the seven "planets," which, in this model, include the Moon and the Sun. Each spherical shell rotates around its center with uniform velocity. If we were to use the model today, say, to predict the apparent position of a star in two hours time, then according to the account I am outlining, we would need to interpret the model in terms of the universe. That is, the model would denote the universe as a whole, and some of its components would denote some of the components of the universe. For example, the star whose position we are interested in would be denoted by one of the stars cast in the sphere of fixed stars. However, in order to interpret the model in terms of the universe, we do not need to assume that the sphere of fixed stars itself or any of the other crystal spheres in the model denotes anything in the universe. So, a user does not need to believe that every object in the model stands for an object in the system.

Second, in order to adopt a certain interpretation of a model in terms of a system, a user does not even need to believe that the objects in the system actually have all of the properties instantiated by the objects that stand for them in the model. For example, one does not need to believe that the string from which a certain pendulum hangs is massless in order to adopt an interpretation of the ideal pendulum according to which the

https://doi.org/10.1086/519478 Published online by Cambridge University Press

string is massless. The knowledgeable user knows perfectly well that, since no real string is massless, the inference, though valid, is not sound and therefore will not actually draw that conclusion about the system from the model. The situation here is analogous to the one we have considered above, in which the knowledgeable user of the map is able to use the old London Underground map successfully in spite of the fact that the map misrepresents some aspects of the London Underground network. Since models often misrepresent some aspect of the system or other, it is usually up to the user's competence, judgment, and background knowledge to use the model successfully in spite of the fact that the model misrepresents certain aspects of the system.

As I have already mentioned, a vehicle can both represent a system and misrepresent some aspects of it. Unlike epistemic representation, *faithful* epistemic representation is a matter of degree. A vehicle does not need to be a completely faithful representation of its target in order to be an epistemic representation of it. In fact, models usually misrepresent aspects of the systems they represent, and it is often up to the user's knowledge and competence to use the models successfully. The user's *background knowledge* will allow her to assess which properties of objects of the model are idealizations or approximations and would lead to unsound inferences about the properties of the corresponding objects in the system. Her *competence* in using the model in representing the system will allow her to judge which valid inferences from the model to the system will be unsound on the basis of her knowledge of how faithfully a certain model represents the various aspects of systems of that kind.

So, if a user does not actually conclude from the ideal pendulum that the string of the actual pendulum is massless, it is not necessarily because their interpretation of the model is more restrictive than the one adopted by someone who does. Rather, it is because they know that inference to be unsound on the basis of their background knowledge and therefore they know that the aspect of the model in question misrepresents the corresponding aspect of the system. This obviously does not mean that the model cannot be an epistemic representation of that system, nor does it mean that it cannot be a faithful epistemic representation of other aspects of the same system. If the latter were the case, the widespread use of idealizations in modeling would be unexplainable.

Similar considerations apply to the model as a whole. The Rutherford model of the atom represents the atom as a small classical planetary system, and the Aristotelian model of the cosmos represents the universe as made up of eight concentric crystal spheres, but the user of these models does not need to believe that the atom *is* a small classical planetary system or that the universe *is* a system of concentric crystal spheres in order to use those models to represent their respective targets.

https://doi.org/10.1086/519478 Published online by Cambridge University Press

7. Interpretation and Surrogative Reasoning. I will now argue that the interpretational conception of epistemic representation allows us to explain why, if a vehicle is an epistemic representation of a certain target, users are able to perform valid surrogative inferences from the vehicle to the target and allows us to tell which inferences from a vehicle to a target are valid. I take it that the fact that the interpretational conception of epistemic representation can explain how users are able to draw valid inferences from a vehicle to a target and what makes some inferences valid and others not are two of the greatest advantages of the interpretational conception. On the inferential conception, the user's ability to perform inferences from a vehicle to a target seems to be a brute fact, which has no deeper explanation. This makes the connection between epistemic representation and valid surrogative reasoning needlessly obscure and the performance of valid surrogative inferences an activity as mysterious and unfathomable as soothsaying or divination. On the interpretational conception, on the other hand, the user's ability to perform pieces of surrogative reasoning not only is not a mysterious skill but it is an activity that is deeply connected to the fact that the vehicle is an epistemic representation of the target for that user.

According to the interpretational conception of epistemic representation, a certain vehicle is an epistemic representation of a certain target (for a certain user) if and only if the user adopts an interpretation of the vehicle in terms of the target. An analytic interpretation underlies the following set of inference rules:

> **Rule 1:** If o_i^V denotes o_i^T according to the interpretation adopted by the user, it is valid for the user to infer that o_i^T is in the target if and only if o_i^V is in the vehicle,
> **Rule 2:** If o_1^V denotes $o_1^T, \ldots, o_n^V$ denotes o_n^V, and ${}^nR_k^V$ denotes ${}^nR_k^T$ according to the interpretation adopted by the user, it is valid for the user to infer that the relation ${}^nR_k^T$ holds among $o_1^T, \ldots, o_n^T$ if and only if ${}^nR_k^V$ holds among $o_1^V, \ldots, o_n^V$,
> **Rule 3:** If, according to the interpretation adopted by the user, o_i^V denotes o_i^T, o_1^V denotes $o_1^T, \ldots, o_n^V$ denotes o_n^V, and ${}^nF_k^V$ denotes ${}^nF_k^T$, it is valid for the user to infer that the value of the function ${}^nF_k^T$ for the arguments $o_1^T, \ldots, o_n^T$ is o_i^T if and only if the value of the function ${}^nF_k^V$ is o_i^V for the arguments $o_1^V, \ldots, o_n^V$.

To illustrate how these rules apply in a concrete situation, suppose that a user adopts the standard interpretation of the London Underground map in terms of the network and that she takes the map to stand for the network. According to Rule 1, from the fact that there is a circle labeled "Holborn" on the map, it is valid for her to infer that there is a station called Holborn on the London Underground network and, from the fact

https://doi.org/10.1086/519478 Published online by Cambridge University Press

that there is no circle or tab labeled "Louvre Rivoli" on the London Underground map, it is valid for her to infer that there is no station called Louvre Rivoli on the London Underground network. According to Rule 2, from the fact that the circle labeled "Holborn" is connected to the tab labeled "Bethnal Green" by a colored line, one can infer that a direct train service operates between Holborn and Bethnal Green station, and from the fact that the circle labeled "Holborn" is not connected to the tab labeled "Highbury & Islington" by any colored line, one can infer that no direct train service operates between Holborn and Highbury & Islington stations.

We are now finally in a position to give a definition of validity for epistemic representation whose interpretations are analytic. If a user adopts an analytic interpretation of the vehicle, then an inference from the vehicle to the target is *valid* (for that user according to that interpretation) if and only if it is in accordance with Rule 1, Rule 2, or Rule 3.

So, if a user is able to perform inferences from a vehicle to a target when the former is an analytically interpreted epistemic representation of the latter, it is because (*a*) a vehicle is an analytically interpreted epistemic representation of the target only when a user adopts an analytic interpretation of it in terms of the target and (*b*) an analytic interpretation of a vehicle in terms of a target underlies a set of rules to perform valid surrogative inferences from the vehicle to the target.

8. A Substantial Account of Scientific Representation. I will now argue that a model represents a system *if* the user adopts an interpretation of the model in terms of the system. In other words, I will argue that interpretation is a sufficient condition for a model to represent a system (as opposed to represent it faithfully). If Suárez was right in thinking that it is not possible to formulate a substantial account of scientific representation, then interpretation could not be a sufficient condition for epistemic (and scientific) representation.

If interpretation was not sufficient for representation, then there would be cases in which a user adopts an interpretation of a model in terms of a system but, nevertheless, the model fails to represent the system. The best way to argue that interpretation is not sufficient for scientific representation is to produce an example in which this is the case.

Suppose that someone "retargets" the Rutherford model of the atom in terms of some arbitrary system, say, a hockey puck sliding on the surface of a frozen pond. Suppose that, according to the new interpretation, the electron in the model denotes the puck and the nucleus denotes the surface of the ice. According to the general interpretation, it would then be possible to infer from the model that, say, the puck is negatively charged and the ice is positively charged, that the puck orbits around the

https://doi.org/10.1086/519478 Published online by Cambridge University Press

ice surface in circular orbits, etc. If the account of scientific representation I propose is correct, it would seem that, under these circumstances, the Rutherford model of the atom is an epistemic representation of the system in question, and, a critic may claim, this is clearly not the case.

This criticism, I think, is based on a conflation of the notions of epistemic representation and partially faithful epistemic representation. It is very likely that, under any standard interpretation, the Rutherford model of the atom only leads to false conclusions about the system in question, or, in any case, we can assume so here. But this is beside the point. Unlike faithful epistemic representation, epistemic representation only requires that it is possible to perform valid inferences from the vehicle to the target and not that these inferences are sound. Thus there are two possible ways to argue that, at least in this case, interpretation is not sufficient for scientific representation. The first is to deny that surrogative reasoning is sufficient for epistemic representation; the second is to deny that interpretation is sufficient for (possible) surrogative reasoning. Since I have already discussed the relation between interpretation and valid surrogative reasoning and since Suárez explicitly denies that surrogative reasoning is sufficient for scientific representation, I will focus on the first option here.

So, on what grounds can one deny that surrogative reasoning is sufficient for representation? A first possible suggestion is that it is insufficient because in the above case *all* the conclusions about the system from the model are false. If this were the case, then a model would be an epistemic representation of a system only if it were a partially faithful epistemic representation of the system. This condition, however, seems to be too strong, as it would rule out that completely unfaithful epistemic representations of a target are nonetheless epistemic representations of that target. For example, suppose that a scientist proposes a bona fide model of a certain system that, upon investigation, turns out to misrepresent every aspect of the system. Even if, gradually, we might discover that all the inferences that are valid according to its standard interpretation are unsound, that model still seems to be an epistemic representation of the system, though an entirely unfaithful one. At no point in time we cease to regard the model as a representation of the system because of its being an unfaithful representation. The Thomson model of the atom is a good historical example of a model that it is very close to being such an entirely unfaithful representation of its target and, nevertheless, it is considered a representation of it.

A second possible suggestion is that the Rutherford model is not an epistemic representation of the puck-on-ice system not because all the inferences from the model to the system are unsound but because the user knows them all to be unsound. If this were the case, then a vehicle would represent a target only if its user does not know all the inferences that

https://doi.org/10.1086/519478 Published online by Cambridge University Press

are valid according to its standard interpretation to be unsound. If, in the above example, the user knew that all inferences from the model to the system are unsound, then the model would not be an epistemic representation of the system.

According to this suggestion, whether something is an epistemic representation of something else is relative to the knowledge the user has of the system. However, in the above example, we have not mentioned anything about the knowledge the user has of the system and, as far as we know, the user could mistakenly believe that at least a few of the inferences that are valid according to the standard interpretation of the Rutherford model are sound. If this were the case, then for that user the model would be an epistemic representation of the target, whereas for us it would not be one. Then there would be nothing intrinsically wrong with claiming that, at least for some users, the Rutherford model could be an epistemic representation of the puck-on-ice system.

Consider again the entirely unfaithful model example. When the model was originally proposed, we did not know that all of the inferences from the model were unsound. Thus, for us, the model was an epistemic representation of the system. However, with time, we gradually find out that all of the inferences from the model to the system were, in fact, unsound. However, there seems to be no reason to think that, when we found out that even the last few inferences were unsound, we ceased to regard the model as a representation of the system. At most, we ceased to regard it as a partially faithful representation of the system, and, as I have argued above, a representation does not need to be partially faithful in order to be a representation.

A third possible suggestion is that the Rutherford model is not an epistemic representation of the puck-on-ice system because no actual user of the model can truly believe that the model allows any sound inference about the target system. This suggestion presupposes that a vehicle can be an epistemic representation of a target only if some user believes the vehicle to be a partially faithful representation of the target. The difference between this suggestion and the first is that, according to this suggestion, the vehicle may turn out to be an entirely unfaithful representation of the target but still be an epistemic representation of it. The difference between this suggestion and the second is that, according to this suggestion, the vehicle can be known, at a later stage, to be an entirely unfaithful representation of the target but it might still be an epistemic representation of it.

Consider again the entirely unfaithful model example. At some point we might have believed that some of the inferences from the model were sound, however this does not need to be the case. Sometimes a model of a system can be put forward as purely hypothetical and conjectural, with-

https://doi.org/10.1086/519478 Published online by Cambridge University Press

out anyone believing that any of the conclusions about the system drawn from the model are going to turn out to be true. The model can be used as a generator of hypotheses about the system, hypotheses whose truth or falsity needs to be empirically investigated. This is often the case when we have little or no idea as to what the internal constitution of a certain system might be, as in the case of atoms in the mid-nineteenth century. Since there seems to be no reason to assume that surrogative reasoning is not sufficient for epistemic representation, I think we can conclude for the moment that it is (unless someone gives us some reason to believe it is not).

So far, I have argued that there is no reason to think that surrogative reasoning is not sufficient for epistemic representation. If one were still to deny that surrogative reasoning is sufficient for epistemic representation, they would either have to assume that there is some "secret ingredient" that is present in the entirely unfaithful model case but missing in the case of the Rutherford model of the atom and the puck-on-ice system. Unless some condition that is met in the first case and not in the second is specified, however, it is difficult to evaluate this claim.

My intuition is that no such secret ingredient can be found. The difference between the two cases, I think, is circumstantial, not substantial. One does not need to have a profound knowledge of physics to realize that the Rutherford model cannot generate any interesting hypotheses about the puck-on-ice system, while the entirely unfaithful model, at the beginning of its career, can be a stimulating hypothesis as to the internal constitution of a certain system. The difference between the two models, however, is not that one represents the system while the other does not—they are both epistemic representations of their target systems. Nor is the difference that one does so faithfully while the other does not—they are both entirely unfaithful.

The difference is only in the role the entirely unfaithful model plays in the wider picture. The exploration of the consequences of the entirely unfaithful model may guide a research program into the constitution of the system under investigation and may be instrumental in the development of better models of that system, while it is clear to us that the Rutherford model will not lead to the discovery of anything we do not already know about the puck-on-ice system and that it is a worse representation of that system than many other simpler and readily available models from classical mechanics.[11]

11. One could probably claim that the Rutherford model of the atom is an uninteresting or unhelpful epistemic representation of the puck-on-ice system to anyone who already supposes that it is an entirely unfaithful representation of that system, but something does not fail to be a representation of a certain target because someone or even everyone

https://doi.org/10.1086/519478 Published online by Cambridge University Press

9. Reconsidering the Inferential Conception. If the arguments in the previous section are correct, then there seems to be no reason to deny that (the possibility of) surrogative reasoning is a necessary and sufficient condition for epistemic representation. So, a substantial account of scientific representation seems possible. In fact, since surrogative reasoning is not only necessary but also sufficient for epistemic representation, then the inferential conception *does* provide us with necessary and sufficient conditions for epistemic representation. If Suárez were to insist that it is not a substantial conception, he should be able to deny that surrogative reasoning is sufficient for epistemic representation—that is, he should produce cases in which a user can perform surrogative inferences from a vehicle to a target without the vehicle being an epistemic representation of the target. As I have argued, I do not think that there is room to do so without blurring the distinction between representation and faithful representation.

If valid surrogative reasoning is already a necessary and sufficient condition for epistemic representation, however, one might legitimately ask why the interpretational conception I have proposed here would be better than a substantial version of Suárez's account, one according to which surrogative reasoning is both a necessary and sufficient condition for epistemic representation such as the one considered in Suárez and Solé (2006).

I think that the main reason to prefer interpretation to surrogative reasoning as a necessary and sufficient condition for epistemic representation is that interpretation is *more fundamental* than surrogative reasoning. It is *in virtue of* their interpretation of the vehicle in terms of the target that users would be able to perform surrogative inferences from the vehicle to the target. While the possibility of surrogative reasoning always accompanies epistemic representation, epistemic representation has conceptual precedence over actual surrogative reasoning. One performs inferences from the London Underground map to the London Underground network *in virtue of* the fact that the map represents the network. The reverse, however, is not true: the map does *not* represent the network *in virtue of* the fact that one uses it to perform inferences about the network. In fact, we would not even attempt to use a piece of glossy paper with colored lines printed on it to find our way around the London Underground network if we did not already regard the former as an epistemic representation of the latter. Surrogative reasoning thus presupposes epistemic representation. Hence, if the map represents the

finds it uninteresting. For example, a map of a railway network that it is no longer in use may be uninteresting to everyone now, but it does not therefore cease to represent that network.

https://doi.org/10.1086/519478 Published online by Cambridge University Press

network, it cannot do so in virtue of its allowing surrogative reasoning (on pain of circularity), but it has to do so in virtue of something else. According to the account defended here, this something else is the fact that the user adopts an interpretation of the map in terms of the network.

The actual performance of surrogative inferences is just a "symptom" that allows us to tell apart cases of epistemic representation from cases of denotation. The actual performance of a surrogative inference from the model to the system reveals that the model is being used as an epistemic representation of the system. In this sense the relation between representation and surrogative reasoning is analogous to that between measles and Koplick spots. Whereas the spots appear if and only if one has measles, this does not mean that one has measles *because* they have Koplick spots.

Suárez would probably agree with my claim that surrogative reasoning is only a symptom of scientific representation. In fact, I suspect that this is his most fundamental reason to think that his account is a deflationary one. The main difference between Suárez's account and mine is that, unlike Suárez, I believe that an account of epistemic representation can and should do more than list the symptoms or surface features of scientific representation.

According to the account I propose, a model represents a target and can be used to perform surrogative inferences about the target *in virtue of* the fact that the user interprets the vehicle in terms of the target. It is the user's interpretation that turns an object into a representation of a certain target.

10. Conclusion: In Search of an Account of Faithful Scientific Representation. In this article I have proposed and defended a substantial account of scientific representation according to which a model is a representation of a certain system in virtue of the fact that a user interprets the model in terms of the system. However, this is only the first step toward a full account of scientific representation. Although in order to have an account of how models represent, it is necessary to have an account of epistemic representation, in and of itself, an account of epistemic representation does not constitute an account of how models represent. To see why, consider a case in which we use the ideal pendulum model to represent the tire swing in the garden. Once we have adopted a certain interpretation, we can infer, say, that the tension of the rope reaches its maximum when the swing is in its rest position. The problem is that, even if this surrogative inference is valid according to the set of rules that our interpretation of the model underlies, the inference may well not be sound. The fact that we have adopted an interpretation only implies that the model represents the system and not that it does so

https://doi.org/10.1086/519478 Published online by Cambridge University Press

faithfully. As I have maintained, the model represents the system faithfully only if all the valid inferences are sound.

But, if this is the case, in virtue of what does the model represent a certain system faithfully? In the pendulum example, we infer that the tension of the rope is highest when the swing is in its rest position from the fact that, in the model, the tension of the rope is highest when the pendulum is in its rest position. But why should what happens in the model tell us anything true about what happens in the system? In order to have an account of how models represent their target system, one's account of representation has to be supplemented with an account of how scientific models represent their target systems faithfully. It is only when we have also a solution to this second problem that we will have a full understanding of how models represent their target systems.

REFERENCES

Bailer-Jones, Daniela M. (2003), "When Scientific Models Represent," *International Studies in Philosophy of Science* 17: 59–74.

French, Steven (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: 1472–1483.

French, Steven, and James Ladyman (1999), "Reinflating the Semantic Approach," *International Studies in the Philosophy of Science* 13: 103–121.

Frigg, Roman (2002), "Models and Representation: Why Structures Are Not Enough," Measurement in Physics and Economics Discussion Paper Series, no. DP MEAS 25/02. London: Centre for the Philosophy of Natural and Social Sciences.

Geiger, Hans, and Ernest Marsden (1909), "On Diffuse Reflection of the α-Particles," *Proceedings of the Royal Society* 82: 495–500.

——— (1999), "Using Models to Represent Reality," in L. Magnani, N. J. Nersessian, and P. Thagard (eds.), *Model-Based Reasoning in Scientific Discovery*. New York: Kluwer/Plenum, 41–57.

——— (2004), "How Models Are Used to Represent Reality," *Philosophy of Science* 71: 742–752.

Hughes, R. I. G. (1997), "Models and Representation," *PSA 1996: Proceedings of the 1996 Biennial Meeting of the Philosophy of Science Association*, vol. 2. East Lansing, MI: Philosophy of Science Association, S325–S336.

Rutherford, Ernst (1911), "The Scattering of α and β Particles by Matter and the Structure of the Atom," *Philosophical Magazine*, ser. 6, vol 21: 669–688.

Suárez, Mauricio (2002), "The Pragmatics of Scientific Representation," CPNSS Discussion Paper Series, no. DP 66/02.

——— (2003), "Scientific Representation: Against Similarity and Isomorphism," *International Studies in the Philosophy of Science* 17: 225–244.

——— (2004), "An Inferential Conception of Scientific Representation," *Philosophy of Science* 71: 767–779.

Suárez, Mauricio, and Albert Solé (2006), "On the Analogy between Cognitive Representation and Truth," *Theoria* 55: 27–36.

Swoyer, Chris (1991), "Structural Representation and Surrogative Reasoning," *Synthese* 87: 449–508.

Teller, Paul (2001), "The Twilight of the Perfect Model Model," *Erkenntnis* 55: 393–415.

https://doi.org/10.1086/519478 Published online by Cambridge University Press