Representing Reality:
The Ontology of Scientific Models and Their Representational Function

Gabriele Contessa

London School of Economics

Thesis submitted to the University of London
for the completion of the degree of a
Doctor of Philosophy

March 2007

To Sherry, Mamma, Papá, and Stella

For their patience and support

I declare that the work presented in this thesis is my own.

Gabriele Contessa

Abstract

Today most philosophers of science believe that models play a central role in science and that one of the main functions of scientific models is to represent systems in the world. Despite much talk of models and representation, however, it is not yet clear what representation in this context amounts to nor what conditions a certain model needs to meet in order to be a representation of a certain system. In this thesis, I address these two questions. First, I will distinguish three senses in which something, a vehicle, can be said to be a representation of something else, a target—which I will call respectively *denotation, epistemic representation,* and *faithful epistemic representation*—and I will argue that the last two senses are the most important in this context. I will then outline a general account of what makes a vehicle an epistemic representation of a certain target for a certain user—which, according to the account I defend, is the fact that a user adopts what I call an interpretation of the vehicle in terms of the target—and of what makes an epistemic representation of a certain target a faithful epistemic representation of it—which, according to the account I defend, is a specific sort of structural similarity between the vehicle and the target.

Contents

Introduction

Most philosophers of science today seem to agree on two basic points. The first is that models play a central role in science. The second is that models are not truth-apt—i.e. they are not capable of being true or false. This has not always been so. Before the so-called syntactic view of theories fell into disgrace in the 1960s, philosophers of science did not take scientific models seriously. The supporters of the syntactic view used to take scientific theories to be collections of sentences or propositions and models to play at most a heuristic role in science. In the 1950s and 1960s, however, the syntactic view came under attack from various fronts and was ultimately repudiated by most of philosophers of science (for an account of the process that led to the demise of the syntactic view and the reasons behind it, see (Suppe 1974b)). After the decline and fall of the syntactic view, the fortune of scientific models in philosophy of science changed dramatically. On the *semantic view* of theories, which originated with the work of Patrick Suppes in the 1960s and can be considered the mainstream view of the structure of scientific theories today (see, for example, (van Fraassen 198) (Giere 1988, Ch. 4) (van Fraassen 1989) (Suppes 1989) (French and Ladyman 1999) (da Costa and French 2003)), scientific theories *are* collections of models, which are the objects that satisfy the axioms of the theory and two sets of axioms are two formulations of the same theory if the same set of models satisfies them both. On the "models-as-mediators" view, which is embraced in one form or another by all the contributors to (Morgan and Morrison 1999) and today is the main alternative to the semantic view, scientific theories are not collections of models. Rather models mediate between abstract theories and concrete systems in the world. On both views, models play a crucial role in science—either scientific theories just consist of models or they play a crucial role in the application of theories to the world. But how do they play this role?

The second point on which most philosophers of science seem to agree despite their respective differences is that models are not truth-apt—they are not capable of being true or false. Scientific models are more like a portrait or a map than like a proposition or a Fregean thought, in that models are not deemed to be capable of being true or false but only more or less faithful or accurate. So, how do models relate to the world? The most popular and promising answer to this question seems to be that they do so in the same way in which a portrait or a map do—which is by representing aspects or portions of the world (see for example, (Cartwirght 1984), (Hughes 1997), (Giere 1999) (Giere 2004)). So for example, the Rutherford model represents the atom and the ideal pendulum model can be used to represent the tire-swing hanging from the tree in the garden.

Despite much talk of models as representations, however, there is widespread disagreement as to what it means to say that a certain model represents a certain system and as to how a certain model represents a certain system. In the literature, the last question is sometimes referred to as "the problem of scientific representation." In the last decade or so, the problem of scientific representation has increasingly attracted the interest of philosophers of science. Unfortunately, this increase in interest in so-called scientific representation has not been accompanied by a comparable increase in our understanding of how models represent systems in the world. This lack of progress, I suspect, is mainly due to the fact that not only it is not clear what the possible solutions to the problem exactly are but it is not even clear what the problem to be solved exactly is.

A casual reader of the literature that deals with the so-called problem of scientific representation might be tempted to believe that there is a well-defined set of worked-out solutions to a clear problem. In fact, I suspect, that this is far from being true. In the literature, it is easy to come across labels such as 'the inferential conception', 'the similarity conception', or 'the structural conception'. However, with the possible exception of the inferential conception, one would be hard pressed to find anywhere in the literature a clear formulation of the views labelled by these expressions that is explicitly endorsed by the alleged supporters of that view. What one can usually find is a number of (mostly casual) remarks rather than anything akin to an account of representation. With a couple of notable exceptions, in this thesis I will be largely unconcerned with what views on scientific representation some author or other holds. This is because too much guesswork and speculation would be involved in figuring out what most authors actually have in mind from what they have explicitly maintained and already too many authors writing on the topic have misinterpreted the scope and content of each other's remarks on the topic. What I will try to do, rather, is to formulate what I take to be the strongest possible version of the views these labels supposedly stand for, the one that I take to best vindicate the intuitions that their supporters seem to have. In most cases, I am not in the position to tell whether the views I will discuss under a certain label reflect the ones actually held by the people who are usually associated with that label but I hope they will be the views that best vindicate their intuitions about representation.

If, as I have claimed, it is far from clear what the possible solutions are, it is even less clear what the problems they are supposed to solve are. In the literature, there seems to be a host of (more or less closely-related) questions that are usually referred to as "the problem of scientific representation". In fact, one of the hypotheses that underlies this thesis is that, properly interpreted, the three conceptions of scientific representation mentioned above are not rival and mutually incompatible (as they are usually taken to be), but they are actually complementary—they should be seen as attempts to account for different notions of representation.

One of the first tasks in tackling the so-called problem of scientific representation will therefore be that of distinguishing the different questions and problems that arise from the use of the notion of representation. In particular, I will distinguish three senses of 'representation' that I take to be particularly important in this context—I call them *denotation*, *epistemic representation* and *faithful epistemic representation.* In this thesis, I will mainly focus on the notions of epistemic representation and faithful epistemic representation.[1] Informally, an

[1] This is not because I do not take the notion of denotation to be important. In fact, denotation plays a crucial role in the account of epistemic representation I defend. If I do not say much about denotation here is rather because, unlike epistemic representation and faithful epistemic

epistemic representation can be thought of as a representation that is used for epistemic purposes (so, beside scientific models, epistemic representations include maps and diagrams but also photographs and portraits). The fact that a user can draw conclusions about the target from an examination of the vehicle is, therefore, a symptom that a vehicle is an epistemic representation of the target (for that user). The fact that some of those conclusions are true is a symptom that the vehicle is a *faithful* epistemic representation of the target. I will consequently identify two corresponding general questions that are often (but not always) conflated in the literature on scientific representation. The first one is 'By virtue of what is a certain object, the *vehicle*, an epistemic representation of another object, the *target*?' The second one is 'By virtue of what is a certain epistemic representation of a certain target *faithful* (insofar as it is)?'

The so-called problem of scientific representation, thus turns out to be neither a special problem nor a single problem. Rather it is an instance of two general problems, which I will call the problem of epistemic representation and the problem of faithful epistemic representation. The main aim of this thesis is to provide general satisfactory solutions to these two problems, a solution which applies to the case of scientific models and real-world systems and, if possible, it also applies to other cases of epistemic representations.

The thesis is divided in three parts. In Part I 'Disentangling Representation,' I will distinguish three relevant uses of the term 'representation.' In Chapter I.1, I will consider a challenge raised by Craig Callender and Jonathan Cohen. According to Callander and Cohen, there is no special problem about scientific representation. Whereas I agree with them about this I think that they think so for the wrong reasons. Callander and Cohen think that there is no special problem of

---

representation, denotation has already received a great deal of attention by philosophers in the analytic tradition and there are already some well-developed accounts of how entities (especially linguistic entities) come to denote other entities. Here I prefer not to commit myself to any specific account of denotation and try to develop accounts of epistemic representation and faithful epistemic representation that are compatible with any account of denotation one might want to embrace.

scientific representation because representation amounts to what I have called denotation—a user just needs to stipulate that the vehicle stands for the target in order for the vehicle to represent the target. I argue that, if representation was purely a matter of stipulation or convention, then any vehicle could represent any target equally well, but there are clearly circumstances in which this is not true. I adapt one of Callender and Cohen's examples to illustrate my point. What Callander and Cohen seem to downplay is that in many cases we want to use the vehicle to learn something about the target—we want to use it as an epistemic representation of the target—and, in those cases, denotation is not a sufficient condition for epistemic representation. In Chapter I.2, I introduce the notions of denotation, epistemic representation, and faithful epistemic representation and I define the last two notions. Then, I discuss some ordinary examples of epistemic representation and faithful epistemic representation and argue that models are epistemic representations of the systems in the world.

In Part II 'Epistemic Representation', I try to formulate a satisfactory answer to the question 'By virtue of what is a certain model an epistemic representation of a certain system?' In Chapter II.1, I consider Mauricio Suárez's inferential conception of scientific representation, which is meant to be an answer to this question. I discuss some of the reasons why I think the inferential conception is unsatisfactory. The two main reasons are the following. First, the inferential conception unwarrantedly stops at what I have called the symptoms of epistemic representation (i.e. the fact that the user is able to perform inferences from the vehicle to the target) and therefore does not really answer the question 'By virtue of what is a certain model an epistemic representation of a certain system?,' which would be answered only by looking at the "causes" of epistemic representation. Second, the inferential conception makes the relation between epistemic representation and surrogative reasoning unnecessarily mysterious. In Chapter II.2, I develop an alternative conception of epistemic representation—the interpretational conception. According to the interpretational concepiton, a vehicle is an epistemic representation of a certain target (for a certain user) only if the user adopts an interpretation of the vehicle in terms of the target. The notion of interpretation plays a crucial role and in determining which inferences from the vehicle to the target are valid and which are not.

In Part III 'Faithful Epistemic Representation', I then move on to develop an account of what makes an epistemic representation of a certain target a faithful one. In Chapter III.1, I argue that the so-called similarity and structural conceptions should be taken to be conceptions of faithful epistemic representation rather than conceptions of epistemic representation. I then consider the similarity conception and argue that, even if the objections that are usually moved against the similarity conception are ineffective, the similarity conception can be effective only if the similarity between the vehicle and the target is an abstract sort of similarity. In Chapter III.2, I develop a version of the structuralist conception of faithful epistemic representation that captures this more abstract notion of similarity—I call this version of the structural conception the *structural similarity conception* of faithful epistemic representation.

If scientific models play a crucial role in science and scientific models relate to the world by representing aspects or portions of it, as most philosophers of science today seem to believe, then understanding how models represent aspects or portions of the world is paramount to understanding how science operates. This thesis lays the foundations to a better understanding of how models represent by accounting for two notions of representation which both play an important role but are too often conflated. Some of the ideas I develop in this thesis build on intuitions and suggestions that can be found elsewhere in the literature. Here, however, these ideas are developed into full-blown accounts of, respectively, epistemic representation and faithful epistemic representation.

For one, a number of authors have suggested that representation has something to do with some structural relation between the vehicle and the target. However, unless it is clearly specified what is the relation between the notion of representation and that of a morphism, how a morphism is supposed to hold between vehicles and targets that are not set-theroretic structure and which morphism needs to hold between a vehicle and a target in order for the vehicle to be a (faithful) epistemic representation of the target, this remains a stimulating idea, but one that can hardly be evaluated. In this thesis, I try to provide clear and definite answers to these in order to turn what can be seen as a stimulating insight about representation into a clear account of a specific notion of representation, whose adequacy can be evaluated.

The notion of (analytic) interpretation plays a central role in this thesis. Most of this thesis was devoted to laying the foundations for a solution for each of these problems. According to the account of epistemic representation that I have defended, the interpretational account of epistemic representation, a vehicle is an epistemic representation of a certain target for a certain user if and only if the user takes the vehicle to denote the target and she adopts an interpretation of the vehicle (in terms of the target). One of the main advantages of the interpretational account that I defend is that it explains the intimate relation between epistemic representation and valid surrogative reasoning. The fact that a user adopts an interpretation of the vehicle in terms of the target (and takes the vehicle to stand for the target) is both that in virtue of what the vehicle is an epistemic representation of the target for her and that in virtue of what she can perform valid inferences from the vehicle to the target. Without the notion of an interpretation (or some analogous notion), the intimate relation between epistemic representation and valid surrogative reasoning would remain unnecessarily mysterious.

The notion of an analytic interpretation plays also a central role in the account of faithful epistemic representation that I develop as it directly contributes to the solution of three crucial problems that have haunted the structuralist conception of (faithful) epistemic representation so far. The first problem is that of applying the notion of a morphism to vehicles and targets that are not set-theoretic structures. As I will argue, the notion of an analytic interpretation provides us with a principled way to reconstruct the vehicle and the target as set-theoretic structures, which I called respectively the relevant structure of the vehicle and the relevant structure of the target.

The second problem is that of determining which morphisms need to obtain between the relevant structure of the vehicle and that of the target in order for the vehicle to be a faithful epistemic representation of the target (to a certain degree) because a morphism may obtain between the relevant structure of the vehicle and that of the target without the first being a faithful epistemic representation of the target (on a certain interpretation of it). The notion of analytic interpretation provides us with a principled way to single out some of the morphisms that may obtain between the relevant structure of the vehicle and that of the target as the intended ones and it is only these morphisms that are relevant to the faithfulness of the epistemic representation in question.

The third problem is that the notion of the faithfulness of an epistemic representation comes in degrees while two structures are either *X*-morphic or they are not. However, the account of faithful epistemic representation that I will develop uses the notions of relvant structure and intended morphism to develop a third crucial notion that of the structural similarity between the vehicle and the target (under a certain interpretation of the former in terms of the latter). Intuitively, the stronger the strongest intended morphism between the relevant structure of the vehicle and that of the target is, the more structurally similar the vehicle and the target are. The central idea that underlies the structural similarity account is that the more structurally similar the vehicle and the target are under a certain interpretation of the former in terms of the latter, the more faithful an epistemic representation of the latter the former is under that interpretation. One of the main merits of the account of faithful epistemic representation that I develop, I think, vindicates the intuitions that underlie two of the main conceptions of representation but avoids the pitfalls that characterize other possible versions of these views.

The account of epistemic representation and that of faithful epistemic representation that I develop and defend in this thesis are more than just complementary—they are deeply interconnected. It is only when one attempts to develop some of the intuitions and ideas that can be found in the literature into a coherent whole that one can see how everything falls into place in the overall picture. In this thesis, I hope to have provided a good initial sketch of that picture.

I. Disentangling Representation

I.1. Is Representation Merely Denotation?

I.1.1. Is There A Special Problem About Scientific Representation?

One of the first questions that face us with regards to scientific representation is whether or not there is any special problem about scientific representation. In other words, do we need to develop a specific conception of how scientific models represent systems in the world? Craig Callender and Jonathan Cohen have recently argued that this question should be answered negatively (Callender and Cohen 2006). Callender and Cohen give the right answer for the wrong reasons.

By arguing that there is no special problem of scientific representation, Callender and Cohen mean to show that the current interest in scientific representation is misplaced because, as they see it, the problem of scientific representation is only an instance of a more general problem of representation to which there is a general solution. In fact, I think, Callender and Cohen simply fail to engage with the sort of questions that those who are interested in the problem of scientific representation are trying to answer. Their failure, however, is instructive in that it will help me to bring into sharper focus what I take to be the two main questions that those who are interested in scientific representation are actually concerned with. In the next chapter, I will distinguish three notions of representation—denotation, epistemic representation, and faithful epistemic representation—and will argue that those interested in scientific representation are concerned with the last two notions of representation rather than the first. So, if there is no special problem about scientific representation, I think, it is because scientific representation is just a variety of what I will call epistemic representation and not because scientific representation is a variety of what I will call denotation as Callender and Cohen seem to think.

I.1.2. REPRESENTATION AND DENOTATION

According to Callender and Cohen, there is a general strategy to reduce a variety of forms of representation to a single fundamental form of representation (most likely, mental representation, which is the representational relation between mental states and what they represent). If successful, this strategy would reduce what would appear as a number of distinct, though interrelated, problems to a single fundamental problem, that of providing an account of mental representation:

> [...] among the many sorts of representational entities (cars, cakes, equations, etc.), the representational status of most of them is derivative from the representational status of a privileged core of representations. The advertised benefit of this [...] approach to representation is that we won't need separate theories to account for artistic, linguistic, representation, and culinary representation; instead, [those who adopt this general approach propose] that all these types of representation can be explained (in a unified way) as deriving from some more fundamental sorts of representations, which are typically taken to be mental states. (Callender and Cohen 2006, p.70)

According to Callender and Cohen, something or other (what I shall call the *vehicle* of the representation) represents something else (what I shall call the *target* of the representation) by virtue of the fact that the speaker intends the vehicle to evoke the target in the mind of an audience. Which mental state a vehicle evokes in an audience and, consequently, which target that vehicle represents for that audience is, according to Callender and Cohen, ultimately a matter of convention or stipulation:

> Can the salt shaker on the dinner table represent Madagascar? Of course it can, so long as you stipulate that the former represents the latter. Then, when your dinner partner asks you what is your

> favorite geographical land mass, you can make the salt shaker salient with the reasonable intention that your doing so will activate in your audience the belief that Madagascar is your favorite geographical land mass (obviously, this works better if your audience is aware of your initial stipulation; otherwise your intentions with respect to your audience are likely to go unfulfilled). [...] On the story we are telling, then, *virtually anything can be stipulated to be a representational vehicle for the representation of virtually anything* [...]. (Callender and Cohen 2006, pp.73–74; *emphasis added*)

If the appropriate conventions are in place, (virtually) anything can be used to represent (virtually) anything else. To use Callender and Cohen's examples, the saltshaker on the table can represent the state of Michigan, if we stipulate so. However, our stipulation is entirely conventional and, had we stipulated that my upturned right hand represented Michigan instead, my right hand would represent Michigan.

If we are to follow Callender and Cohen, thus the prototype of (derivative) representation is the relation between referring expressions and their referents. In Italian, the word 'gatto' is used to refer to cats and 'cane' to refer to dogs, but this choice is entirely arbitrary: there seems to be no intrinsic reason why it should be so and not, say, the other way around. (By saying that there is no *intrinsic* reason, I mean that there is no intrinsic property of the word 'gatto' that makes it preferable to use it to designate cats rather than dogs).

If we were to follow Callender and Cohen's strategy, then we would have to maintain that the Rutherford model of the atom represents the atom by virtue of the fact that Rutherford intends it to stand for the atom and that, to fulfil his intentions, Rutherford has tacitly stipulated with his audience that it does so. On Callender and Cohen's view, it would seem that Rutherford could as well have intended that the ideal pendulum, the Thompson model of the atom, or even the paper weight on his desk represented the atom.

To avoid the most implausible consequences of their view, Callender and Cohen have to concede:

> [...] it should be clear that the constraints ruling out these choices of would-be representational vehicles are pragmatic in character: they are driven by the needs of the representation users, rather than by essential features of the artifacts themselves. (Callender and Cohen 2006, p. 76)

So, even if *in principle* it would have been possible for Rutherford to use anything to represent the atom, *in practice* some vehicles are *more convenient* than others. This, however, according to Callender and Cohen, does not mean that, in principle, a saltshaker could have not been used to represent the atom as well. Whether this line of defence is convincing crucially depends on what 'more convenient' means in this context.

Obviously, 'more convenient' does not mean that Rutherford found it easier to conceive of a new model of the atom rather than choosing any object from his desk and stipulating that that object represented the atom, or to pick any other model available at the time, say, the ideal pendulum, and stipulate that that represented the atom. Had any of these objects been able to serve Rutherford's purposes equally well, it would have been certainly easier for Rutherford to use one of them rather than conceiving of a new model of the atom. Thus 'more convenient' does not mean easier.

The sense in which 'more convenient' has to be construed if it is to serve Callender and Cohen's purposes clearly emerges from one of their examples. If we want to show someone where a place is in Michigan, they claim, it would be more convenient to represent Michigan by an upturned right hand than by a saltshaker. This, according to Callender and Cohen, is because:

> [...] the geometric similarity between the upturned human right hands and the geography of Michigan make the former a particularly useful way of representing relative locations in

> Michigan, and it normally would be foolish (but not impossible!) to use an upturned left hand for this purpose since a more easily interpreted representational vehicle is typically available. (Callender and Cohen 2006, p.76)

By conceding this much, however, Callender and Cohen reveal how unsatisfactory their strategy in fact is. In this case as in many others, our choice of a vehicle is not completely conventional. As Dominic Lopes (1996 pp.132–133), following David Lewis' analysis of convention (Lewis 1969, p.76), notes against Nelson Goodman's view of pictorial representation, a choice is conventional (with respect to a certain set of alternatives) only if those who make it have no intrinsic reason to prefer one of the available alternatives over the others. For example, the choice of the word 'gatto' to designate cats and 'cane' to designate dogs or that of driving on the right-hand side of the road rather than on the left-hand side are conventional (with respect to a certain set of alternatives) insofar as there is no reason to prefer any of the other alternatives to them. To the extent to which we have some intrinsic reasons for preferring one of the options over the others, however, our choice is no longer entirely conventional.

If we want to show someone where a certain place in Michigan is, our preferences are clear: we would prefer a map over an upturned right hand and an upturned right hand over the saltshaker. Callender and Cohen not only acknowledge our preferences but go as far as suggesting that the reason why we prefer an upturned right hand to a saltshaker as a representation of the geography of Michigan is the "geometric similarity" between the hand and Michigan. To maintain that our choice of the map rather than the saltshaker as a representation of the geography of Michigan is conventional because, though foolish, it would not be impossible to choose the saltshaker, as Callender and Cohen suggest in the above passage, is analogous to maintaining that, if we are offered the choice between a gift of £2 and a gift of £200,000 with all else equal, our choice between the two options would be conventional because, though foolish, it would not be impossible to choose to win £2 rather than £200,000.

Although in the case of the map and the salt shaker our preferences are clear, the reasons that underlie them are difficult to articulate. Why is a saltshaker less preferable as a candidate for a representation of the state of Michigan than an upturned right hand and an upturned right hand less preferable than a map? Why did Rutherford go through the trouble of conceiving a new model of the atom rather than simply using one of the objects on his desk? Much work needs to be done to answer these questions adequately. In the next chapter, I will lay the foundations for this work by distinguishing three notions of representation—denotation, epistemic representation and faithful epistemic representation. It is only in Chapter II.2, where I develop and defend my account of epistemic representation, that I will be able to suggest what I take to be a satisfactory answer to this question.

I.2. Denotation, Epistemic Representation, and Faithful Epistemic Representation

I.2.1. Epistemic Representation and Surrogative Reasoning

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 two different senses of 'representation'. In a first sense, both the logo of the London Underground and a map of the London Underground can be taken to represent the London Underground network. In the terminology I adopt here, we can say that they both *denote* the network (I will say more about denotation in Section I.2.2 below). The map of the London Underground, however, 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 network in this stronger sense that, to use the terminology introduced by Chris Swoyer (1991), someone can perform *surrogative inferences* from the map to the network—that is one can infer conclusions about the network from a consideration of the map. If *A* and *B* are two distinct objects, an inference from *A* to *B* is *surrogative* if and only if the premise of the inference is a proposition about *A* and the conclusion of the inference is a proposition about *B*. In our example, the map and the network are clearly two distinct objects. One is a piece of glossy paper on which coloured lines, circles and names are printed; the other is an intricate system of, among other things, trains, tunnels, rails and platforms. Yet, the ordinary users of the map (and of the network) frequently perform surrogative inferences from one object, the map, to the other, the network. For example, from 'the circle labelled Holborn is connected to the circle labelled Liverpool street by a red line' (which expresses a proposition about the network) users infer 'Central Line trains operate between Holborn and Liverpool street station' (which expresses a proposition about the network). The same does not apply to the London Underground logo. Users of the London Underground network do not usually use the London Underground logo to perform surrogative inferences from it to the network and there seems to be no obvious way to do so. The logo may denote the network but it is not an epistemic representation of it (I will say more about this in Section I.2.3 below).[2]

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 a "symptom" of the fact that, for that user, that vehicle is an epistemic representation of that target, a symptom that allows one 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 or other 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 surrogative inferences if the occasion had arisen.

It is also important to note that it is not necessary that the conclusions the user draws (or would draw) about the target are true (or approximately true) in order for the vehicle to be an epistemic representation of the target. In other words, if a user is able to perform (valid) inferences from a certain vehicle to a certain target, the

[2] Note that I am only claiming that the logo *is not* an epistemic representation of the network not that it *could not* become one if the appropriate conditions obtained. I take it that an account of epistemic representation is minimally one that specifies what conditions need to hold for something to be an epistemic representation of something else.

vehicle is an epistemic representation of that target for that user, independently of whether or not the conclusions are true of the target.

Even if I have not yet introduced the full conceptual apparatus that is needed to define precisely when a surrogative inference is valid (I will do so only in Section II.2.6 below), it is useful at this point to distinguish between valid and sound surrogative inferences. For the moment, it is sufficient to note that, whereas a surrogative inference is sound if and only if it is valid and its conclusion is true (or approximately true),[3] the conclusion of an inference does not need to be true in order for the inference to be valid. I trust that the distinction between valid and sound surrogative inferences is clear enough for the present purposes—a valid inference is sound only if its conclusion is true (or approximately true). However, since the notion of a valid surrogative inference will play a role in this chapter and the next, I will characterize it informally here.

The intuitive idea behind the notion of valid surrogative inference is that a surrogative inference is valid only if it is in accordance with a systematic set of rules. For example, according to the set of rules standardly associated with the London Underground map, it is valid to infer that Holborn and Bethnal Green stations are connected by Central Line train form the fact that the circles marked 'Holborn' and 'Bethnal Green' are connected by a red line. According to the same set of rules, however, from the fact that the circles marked 'Holborn' and 'Bethnal Green' are 3.4 inches away it is not valid to infer that, say, the distance between Holborn and Bethnal Green station is 3.4 miles (or anything else about the London Underground network for what matters). That inference is simply not allowed by the standard rules associated with the map. However, it is in principle possible to devise a non-standard set of rules, according to which the latter inference is valid and the former is not.

---

[3] Let me note that 'sound' here is not used in the same sense in which it is used in logic where an inference is sound if it is valid and its premises are true. I will say more about what I mean by approximately true in Section I.2.5.

By specifying that the set of rules be systematic, I intend to stress that the result of applying a certain rule does not depend on who applies it or the circumstances in which she applies it but only on the way the vehicle is and what the rule says. A systematic set of rules, for example, cannot include rules such as: 'If the circles marked 'Holborn' and 'Bethnal Green' are connected by a red line, conclude the first thing that goes through your mind' because such a rule would give different results when different people apply it or when the same people apples at different times.

Unless otherwise specified, when talking about a valid inference from a certain vehicle to a certain target, I will do so by assuming that there is a set of rules standardly associated with that vehicle and that the inference in question is one of those that are in accordance with that set of rules. Whether there are any such rules, what there rules exactly are and where they come from are all crucial questions, which, however, I shall address only later on in Section II.2.6 below after having introduced my account of epistemic representation and the crucial notion of interpretation.

For the moment, however, it is possible to formulate the following definition:

(1) A vehicle is an *epistemic representation* of a certain target for a certain user if and only if:

[1.1] the user is able to perform valid (though not necessarily sound) surrogative inferences from the vehicle to the target.

I will call condition [1.1] *valid surrogative reasoning.* Definition (1) thus says that valid surrogative reasoning is a necessary and sufficient condition for epistemic representation.

Let me note an important feature of the notion of epistemic representation. According to (1), 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).[4] 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.[5]

The notion of epistemic representation is primarily a technical notion. As with 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 turn out to be epistemic representations for us. Portraits, photographs, maps, graphs, and a large number of other representational devices usually seem to allow us to draw (valid) inferences to

---

[4] 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. Suárez 2002 and 2003, Frigg 2002, 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 accounts 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, in the past, the likes of Giere and French may have given the impression that they conceived of representation as a dyadic relation, I do not think that this is their considered view. 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, p.743).

[5] 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 does not imply that it is an epistemic representation in and of itself.

their targets and, if, as such, they are epistemic representations of their targets (for us) according to definition (1). For example, according to (1), if we are able to draw (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 more or less faithful (epistemic) representation of that target. Let me illustrate this point with an example. If we show a friend a portrait that depicts a 17th century nobleman and ask her to describe the person represented by it, she will probably be able to infer from the portrait a description of the person portrayed. In doing that, our friend performs a number of inferences from the portrait to the person it portrays. But why should we assume that the description inferred from the portrait is true? For all we know, the portrait could be a erroneous representation of its subject (the painter may have never actually seen the person portrayed or may be mistaken about them), it could be a mendacious representation of its subject (the painter intended to mislead the viewers about the appearance of its subject), or it could be an ironic representation of its subject (who is in fact a contemporary whose manners resemble those of a 17th century gentleman).

Suppose, for example, that the painter meant to mislead the viewers about the appearance of the person portrayed, say, the Duke of Edinburgh. The portrait can deceive the viewers about the appearance of the Duke of Edinburgh only if, for them, it is an epistemic representation of Duke of Edinburgh—i.e. only if, from it, the users are going to draw conclusions about the appearance of the Duke of Edinburgh not someone else. However, if the viewers are deceived by the portrait, at least some of the valid inferences from the portrait to the Duke must not be sound. For this reason, besides notion of epistemic representation we need to introduce the further notion of faithful epistemic representation, which I will do in Section I.2.6.

I.2.2. Some Remarks About Denotation

Denotation is among the most widely discussed notions in philosophy and it is well beyond the scope of this thesis to give a full account of this notion. Nevertheless, since this notion plays an important role throughout this thesis, a few remarks about my use of the notion seems to be in order here.

The notion of denotation is usually discussed within the context of philosophy of language. There denotation is (minimally) construed as a relation that holds between certain kinds of linguistic expression—I will call them *denoting expressions*—and objects (in the broadest sense of the word).[6] Proper names are the prototypical example of denoting expressions. For example, 'Napoleon' denotes Napoleon.

In the analytic philosophy of art, however, the notion of denotation has a broader use. On this broader use, none of the relata of the denotation relation needs to be a linguistic expression. Any two objects (in the broadest sense of the term 'object') can be (among) the *relata* of the denotation relation. As far as I can see, this use of the term 'denotation' in philosophy of art is consistent with the narrower use that seems to be prevalent in philosophy of language if we only assume that the objects that serve as linguistic expressions are only some of the objects that can be used to denote other objects. In this thesis, I use 'denotation' in the broader sense in which it is used in philosophy of art.

But what is denotation? What does it mean to say that something denotes something else? I will not try to answer these questions here. Rather I will take denotation to be a primitive notion This does not mean that philosophers or scientists cannot tell us anything deeper about denotation (e.g. what is it about the

---

[6] I say 'minimally' because I take it that most philosophers of language do not conceive of the denotation relation as a dyadic relation.

human mind or brain that allows us to use objects to denote other objects). It only means that, even if there is something deeper to be said about denotation, it is not likely to affect what I will say here.

Here, I will assume that a vehicle denotes a certain target for a certain user if and only if the user takes the vehicle to denote the target. As far as I can see, I do not need to commit myself to any specific account of what it means for a user to take an object to denote another object. I take it that it is uncontroversial that, as human beings, we happen to be able to use some objects to denote (or stand for) other objects. If one were to deny that we have this ability, they would seem to have troubles to explain even our most basic linguistic and symbolic practices. If there is any controversy about denotation, it concerns *how* some things get to denote other things (for us) not about the fact *that* some things get to denote other things (for us). However, nothing in what I will say here depends on *how* some things get to denote other things for us. It only depends on the fact that some things do denote other things for us, which, as I said, I take to be a rather uncontroversial assumption.

I.2.3. Denotation and Epistemic Representation

Once we distinguish between, denotation and epistemic representation, it is easier to see where exactly I disagree with Callender and Cohen. I think that Callender and Cohen are right in believing that there is no special problem of scientific representation—it is their reasons for believing so that are wrong. Callender and Cohen think that there is no special problem about scientific representation because scientific representation is a case of denotation. I believe that there is no special problem about scientific representation because it is a case of epistemic representation.

In particular, it may well be the case that in principle anything can *denote* anything else for a user if the right conditions are in place, but this does not mean that denotation is a sufficient condition for epistemic representation. The London Underground logo, the words 'the Tube', and the London Underground map may all denote the London Underground network (for us), but only the last is an epistemic representation of it for us—only the map is used by us to perform pieces of surrogative reasoning from it to the London Underground network.

It is important to note that, here, I am not denying that the London Underground logo and the words 'the Tube' *could be* epistemic representations of the London Underground network *if the appropriate conditions obtained.* What I am denying is that, even if they denote the London Underground network, they *are* epistemic representations of it. I take this to show that denotation is not a sufficient condition for epistemic representation.

I.2.4. Representational Content, Scope, and Background Knowledge

I will now introduce some further notions that will be useful in what follows.

(2) Two epistemic representations of the same target offer *conflicting representations* of (an aspect of) the target if and only if:

[2.1] From the one it is valid to infer a certain proposition and from the other is valid to infer the negation of that proposition.

So, for example, an old 1930s map of London Underground and the new map offer conflicting representations of (an aspect of) the London Underground network because, among other things, from the one it is valid to infer that there is no direct train service between Euston and Oxford Street stations, while, from the other, it is valid to infer that there is a direct train service between those stations.

(3) The *representational content* of an epistemic representation is the set of propositions that is valid to infer from the vehicle (in accordance with a given set of rules).

So, for example, since, according to the rules standardly associated with the London Underground map, it is valid to infer from it that Holborn and Bethnal Green stations are connected by Central Line trains, the proposition expressed by 'Holborn and Bethnal Green stations are connected by Central Line trains' is part of the representational content of the map (given the standard set of rules ordinarily associated with it).

It is important to emphasize that a vehicle in and of itself does not have a representational content only epistemic representations do. It is only in relation to a set of rules according to which some surrogative inferences are valid and some are not that the epistemic representation has a representational scope. Since we ordinarily use the same term (e.g. 'map') to refer to an epistemic representation and its vehicle, it is easy to get confused about this. However, in any ambiguous cases, context should help the reader to determine whether terms such as 'map' refer to an epistemic representation or its vehicle. For example, whenever I talk of the representational content of a map, I intend 'map' to refer to an epistemic representation not to the material object that is its vehicle.

(4) Two epistemic representations of the same target have the same *scope* if and only if:

[4.1] for every proposition *p*, *p* is part of the representational content of the one if and only if *p* or its negation is part of the representational content of the other.

(5) Of two epistemic representations of the same target, *A* and *B*, *A* has wider scope than *B* if and only if:

[5.1] for every proposition *p*, if *p* is part of the representational content of *B*, *p* or its negation is part of the representational content of *A*.

[5.2] for some proposition *p*, *p* is part of the representational content of *A* and neither *p* nor its negation is part of the representational content of *B*.

(6) Of two epistemic representations of the same target, *A* and *B*, *B* has narrower scope than *A* if and only if:

[6.1] *A* has wider scope than *B*.

It follows from (2) and (4) that two epistemic representations can have the same scope even if they offer conflicting representations of their target. So, for example, if two maps of Venice have exactly the same representational content except for the fact that from one it is valid to infer that there is a bridge over a certain canal while from the other it is possible to infer that there is no bridge over that canal the two maps have the same scope but they offer conflicting representations of (that aspect of) Venice.

It is important to note that definitions (4) and (5) imply that the relation 'having the same or broader scope than' is a partial order. In other words, it is possible that neither of two epistemic representations of the same target has the same or broader scope than the other. This seems to be intuitively correct if one considers that there are epistemic representations of the same target whose scopes are largely different and therefore hardly comparable. For example, if a view of Venice and a map of Venice have largely different scopes, there seems to be no point in trying to determine which of them has wider scope.

I.2.5. ASIDE: APPROXIMATE TRUTH AND CLOSENESS TO THE TRUTH

In the previous section, I have defined a sound surrogative inference as one that is valid and whose conclusion is true or approximately true. In this section, I will clarify my use of this expression (on the understanding that a philosophical account of the notion of approximate truth is well beyond the scope of this thesis).

As I use it here, the expression 'approximately true' applies to sentences or propositions. A sentence or proposition is approximately true if it is "close enough" to the truth, where in most cases what counts as close enough depends on the user and the circumstances. For example, assume that (the proposition expressed by) 'The room is 6 feet wide' is strictly speaking false because the room in question is actually 5'11" wide. Though strictly speaking false, it is clear that there are circumstances in which we would not consider (the proposition expressed by) 'The room is 6 feet wide' false. For example, if we are describing someone's house to a friend, it is unlikely that our friend would consider our description of the house false upon discovering that the room is actually 5'11" wide. *For us* and *in those circumstances*, the (proposition expressed by the) sentence 'the room is 6 feet wide' would be approximately true. In different circumstances, however, the same sentence would not be approximately true for us. For example, if our friend wants to know whether a 6-feet-wide bookshelf will fit in one of the rooms we have just described, the sentence would no longer be approximately true because our friend needs to know if the room is wider than 6 feet. (If the room was actually 6'1', 'The room is 6 feet wide' would probably still be approximately true for us in those circumstances.)

However, there is also a sense in which 'The room is 6' wide' is *closer to the truth* (about the width of the 5'11" wide room) than 'The room is 6'1" wide'. By this I only mean that, for *all* circumstances and *all* users, if the second sentence was approximately true, the first would be approximately true, but there are circumstances in which the first is approximately true but the second is not. (For example, if our friend's bookshelf was exactly 6'1" wide the first sentence would still be approximately true for us while the second would be not. From the first we could find out that the bookshelf doesn't fit on that wall, from the other we would not.)

I.2.6. FAITHFUL EPISTEMIC REPRESENTATION

Consider again the old map of the London Underground and the new map of the London Underground I mentioned at the beginning of the previous section. Both represent the London Underground network in the sense that one can perform surrogative inferences from either map to the network. However, the two maps offer two conflicting representations of (some aspects of) the London Underground network—one can validly draw from one some conclusions that are the negation of conclusions that can be validly drawn from the other. As I have mentioned, 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 those two stations.

Whereas the surrogative inferences from the new map are all sound (or at least so I assume here), some of the inferences from the old map to today's network are not because, in the meantime, there have been significant changes to the London Underground network (I assume that the reverse would be true of the 1930's London Underground network). In this sense, we could say that the old map *misrepresents* today's network (to some degree). *Prima facie*, it could seem inconsistent to claim that the old map both represents today's network and misrepresents it. However, the inconsistency disappears once we realize that the fact that the map misrepresents the network not only does not imply that the map does not represent the network but actually implies that it represents it.

To avoid this kind of confusion, it is wise to introduce the following set of definitions:

(7) A vehicle is a *completely faithful epistemic representation* of a certain target if and only if:

[7.1] it is an epistemic representation of the target and

[7.2] all of the valid inferences from it to the target are sound.[7]

(8) A vehicle is a *partially faithful epistemic representation* of a certain target if and only if:

[8.1] it is an epistemic representation of the target and

[8.2] some of the valid inferences from it to the target are sound.

[7] It is important to note that, in order to be a completely faithful epistemic representation of its target, a representation does not need to be a perfect replica of its target. The new London Underground map, for example, is a completely faithful epistemic representation of today's network because, from it, one can only draw true conclusions about today's network (or, at least, so I assume here). This, however, does not mean that the map is 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 (e.g. the internal structure of the stations or the spatial relations among them).

(9) A vehicle is a *completely unfaithful epistemic representation* of a certain target if and only if:

[9.1] the vehicle is an epistemic representation of the target, and

[9.2] none of the valid inferences from the vehicle to the target are sound.

If only we assume that by 'misrepresents' we mean 'is not a completely faithful epistemic representation of' and by 'represents' we mean 'is an epistemic representation of the network,' then it becomes apparent that not only the fact that the map *mis*represents today's network does not imply that the map does not represent the network it but it actually implies that it does represent it. In other words, a vehicle cannot be an unfaithful epistemic representation of a certain target unless it is an epistemic representation of that target.

The notions defined in (7), (8) and (9) concern what I will call the *overall* faithfulness of a certain epistemic representation. However, the overall faithfulness of an epistemic representation should be distinguished by its specific faithfulness. For example, if we are only interested in finding out whether or not there is a direct train connection between Holborn station and Liverpool Street station, the old London Underground map is as faithful a representation of the London Underground network as the new London Underground map *for this specific purpose*, even if *overall* the latter is a more faithful epistemic representation of it than the former.

(10) A vehicle is a *specifically faithful epistemic representation* of a certain target for a certain user and a certain purpose if and only if:

[10.1] it is an epistemic representation of the target

[10.2] and

[10.3] the specific conclusion in which the user is interested is true.

Unlike epistemic representation, faithful epistemic representation is a matter of degree. An epistemic representation can be a more or less faithful epistemic representation of its target. The same vehicle can be a faithful epistemic 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 today's network. However, from the old map, it is also possible to draw a number of true conclusions about today's network. In general, an epistemic representation of a certain target is faithful only insofar as the inferences from the vehicle to the target are sound.[8]

Two faithful epistemic representations, *A* and *B*, of the same target *T* can be equally faithful epistemic representations of it or one can be a more faithful epistemic representation than the other. It is therefore important to introduce the following two definitions.

(11) Two epistemic representations, *A* and *B*, are *equally faithful epistemic representations* of *T* if and only if:

[11.1] *A* and *B* have the same scope

[11.2] the set of true conclusions validly drawn from *A* to *T* and the set of true conclusions validly drawn from *B* to *T* coincide and

[11.3] None of the false conclusions validly drawn from *A* or *B* to T is closer to the truth than the corresponding conclusion from *B* or *A* to *T*.

(12) Of two epistemic representations of *T*, *A* and *B*, *A* is a *more faithful* than *B* if:

[8] The distinction between epistemic representation and faithful epistemic representation is analogous to a distinction drawn by Mauricio Suárez (Suárez 2002 and Suárez 2004), who claims that we should distinguish between representation and 'accurate, true and complete representation' (Suárez 2004, p.767). The importance of such distinctions, I think, can hardly be overestimated. The fact that 'representation' is used to both refer to what I call epistemic representation and faithful epistemic representation is the source of many of the problems connected with the notion of representation. Much misunderstanding could be easily avoided by carefully distinguishing between different senses of representation.

[12.1] *A* and *B* have the same scope and

[12.2] :

[12.2.1] the set of true conclusions about *T* that it is valid to draw from *B* is a proper subset of the set of true conclusions that it is valid to draw from *A* or

[12.2.2] :

[12.2.2.1] some of the conclusions validly drawn from A to T are closer to the truth than those from B to T and

[12.2.2.2] none of the conclusions validly drawn from *B* to *T* are closer to the truth than those from *A* to *T*.

It is important to note that the conditions provided by (12) are sufficient but not necessary conditions for *A* to be a more faithful epistemic representation of *T* than *B*. There may well be other ways for *A* to be a more faithful epistemic representation of *T* than *B* that are not covered by (12). For example, if from *B* it is valid to infer only one true conclusion about (a certain aspect of) *T* and from *A* it is valid to infer a very large number of true conclusions about other aspects of *T* but a false conclusion about that particular aspect of *T*, I think that we would be likely to regard *A* as a more faithful epistemic representation of *T* than *B* even if *A* does not satisfy the conditions provided by (12). However, I do not think that these cases are as uncontroversial as the cases (12) covers. It seems to be at best unclear how the unfaithfulness of a representation relates to that of the other. The virtue of (12), I think, is that it covers all cases that are uncontroversial while leaving open the possibility that some less uncontroversial cases may still be cases of one epistemic representation of a certain target being more faithful than another.

Even if it was possible to provide a set of necessary and sufficient conditions for an epistemic representation of a target to be more faithful than another, the relation 'being an equally or more faithful epistemic representation of *T* than' would not be a partial order. That is, there are epistemic representations of a certain target such that neither is an equally or more faithful epistemic representation of the target than the other. This follows from the fact that among the conditions for being an equally faithful epistemic representation of a certain target than another is that both epistemic representations have the same scope and not all representations of the same target have the same scope (in the abovementioned case of the view of Venice and the map of Venice for example we have reasons to believe that the two representations have largely different scopes).

Some may think that the requirement that A and B have the same scope in (11) is too stringent. However, I think that not including that requirement would give us some counterintuitive consequences. For example, we could have two epistemic representations of a certain target that are equally faithful even if one is completely faithful and the other is not (if the first has narrower scope than the second). I think that intuitively we would think that a completely faithful epistemic representation of a certain target cannot be equally faithful to one that is not completely faithful as, by definition, an epistemic representation that is not completely faithful is partially unfaithful.

I.2.7. SCIENTIFIC MODELS AS EPISTEMIC REPRESENTATIONS

How does all this relate to scientific models? Scientific models, I claim, are epistemic representations of certain systems in the world (for their users) and, as such, they can be used to perform surrogative inferences from the model to the systems in question. In this section, I illustrate this claim by means of three emblematic examples of scientific models: the inclined plane model (I.2.7.1) and the Thomson and the Rutherford models of the atom (I.2.7.2).

I.2.7.1 The Inclined Plane Model

Suppose that we want to ensure that a soap-box derby is safe. More specifically, we want to make sure that, when the cars reach the finish line at the foot of the hill, their velocity does not exceed a specific velocity that we deem safe. To perform this task, we may use a very simple model from classical mechanics: the frictionless inclined plane.[9] In the model, a block lies at the top of a frictionless inclined plane. The only forces acting on the block are the gravitational force and the normal force acting perpendicular to the plane. At the top of the plane, the kinetic energy of the block is 0 ($KE_i=0$) and its gravitational potential energy is $U=mgh$, where m is the mass of the block, g is the gravitational acceleration and h is the height between the top and the bottom of the plane. The mechanical energy of the system, E, is thus given by $E=KE_i+U=0+mg=mgh$. Since the energy in the system is conserved, when the block reaches the end of the slope, all the potential energy will be converted into kinetic energy. The kinetic energy of the block at the end of the slope will be, therefore, $KE_f=1/2\ mv^2=mgh$. Solving for v, we can determine that the velocity of the block at the bottom of the slope will be $v_f=(2gh)^{1/2}$. The final velocity of the block, therefore, depends only on h and g.

By plugging in determinate values for g and h, we obtain a determinate value for the velocity of the block at the bottom of the slope. From physics textbooks, we know that 9.8m/s^2 is a very good approximation of the gravitational acceleration experienced by a body near the surface of the earth, such as the cars in the soap-box derby, and we can set h so that it is the same as the difference in height between the start-line of the soap box-derby and the foot of the hill, say, 10 metres. Given these inputs, the velocity of the block at the bottom of the slope will be approximately 14 m/s or 50.4 km/h.

From the inclined plane model, we can therefore infer that the velocity of the cars at the bottom of the hill is going to be 50.4 km/h. A knowledgeable user however will take this conclusion with a pinch of salt, as she knows that the inclined plane model cannot possibly be a completely faithful epistemic representation of the soap-box derby. One of the reasons to think so is that some of the factors that affect the final velocity of the cars have no counterpart in the model. The knowledgeable user would probably say that the final velocity of the cars will be approximately 50 km/h and, since the most relevant of these factors

[9] I owe this example to Bernard Nickel.

contribute to the deceleration of the cars, she would probably say that the velocity of the cars is likely to be less than 50 km/h.

Here it is important to distinguish what the user infers from the model from what her back-ground knowledge tells her about the conclusion drawn from the model. What she infers from the model about the velocity of the cars is not different from what any other competent user of the model would normally infer from it; however her background-knowledge suggests to her that the conclusion drawn from the model cannot be literally true and is, at best, "approximately true"—that is the value of the velocity of the cars will be close enough to that of the velocity of the box in the model.

I.2.7.2 The Thomson and the Rutherford Models of the Atom

Consider now a more complicated case. The Rutherford model of the atom was originally proposed by Ernest Rutherford (1911) in order to account for the phenomenon now known as Rutherford scattering. In a series of experiments in 1909, Hans Geiger and Ernest Marsden 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 Rutherford's experiment was 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. Simple calculation shows that, for each atom an alpha particle crosses, an alpha-particle would pick up a total sideways velocity of approximately 6750 metres per second—only a few tenthousandths of the particle's forward velocity, which is approximately 1.6 x 10^7 meters per second (see Fowlers 1997).

It is worth noting that here we are using a model, the Thomson model of the atom, as a building block of a larger model, a model of Geiger and Masden's experiments. In the latter model, alpha particles are represented as very heavy, positively charged particles and each of the atoms in the golden foil is represented as the atom in the Thomson model. From this larger model, we can infer that Rutherford scattering would never occur. However, since Geiger and Masden's 1909 experiments show that the phenomenon actually occurs, the model of the experiment based on the Thomson model of the atom is not a completely faithful epistemic representation of the experimental situation as it leads to a false conclusion about the experimental results. As Rutherford quickly realized, the experiments showed that the Thomson model was not a faithful epistemic representation of the atom.

The results of the Geiger and Marsden experiments suggested to Rutherford a different model of the atom. In the Rutherford model, all the positive charge of the atom and almost all of its mass is concentrated in the nucleus, whose radius is one-hundredth of that of the atom, and the rest of the volume of 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 little to scatter high-momentum alpha-particles, most alpha-particles will not be deflected. Unlike the Thomson model of the atom, thus, from the Rutherford model we can soundly infer that Rutherford scattering will occur. From the model, we can also infer that the scattering of one in 20,000 alpha-particles at large angles is caused by the electromagnetic repulsion exerted by the atomic nuclei on those alpha particles that go close enough to one of them. If this inference is sound, the Rutherford model faithfully represents this aspect of the behaviour of the atom and explains Rutherford scattering.

When we say that the Thomson and the Rutherford models of the atom represents one of the atoms in the golden foil in Geiger and Marsden's experiments, we are not merely saying that they denote the gold atom, like the letters 'Au' on the periodic table do. Rather, we are saying that they are epistemic representations of the atom—in the sense, that both can be used to draw conclusions about certain aspects of atoms. As I have already mentioned, the two models are not equally faithful epistemic representations of the atom. One of the conclusions about the atom that can be drawn from the Thomson model of the atom (according to its standard interpretation) has been proven false by Geiger and Marsden's experiments.

Only a couple of years after Rutherford proposed his model of the atom, in 1913, Niels Bohr published a pioneering paper in two parts titled 'On the Constitution of Atoms and Molecules' (Bohr 1913). Among other things, in the paper, Bohr pointed out that the Rutherford model of the atom is highly unstable. According to classical electrodynamics, any accelerated charge radiates energy. Therefore, the orbiting electrons in the Rutherford model would rapidly collapse into the nucleus. The Rutherford model was not a completely faithful epistemic representation either as from it one could draw the conclusions that atoms are much more unstable than they actually are.

It is difficult to assess to what extent this result was new to Rutherford. On the one hand, nothing in the Rutherford 1911 article suggests that Rutherford *knew* that the atom in his model was extremely unstable. This hypothesis is also supported by the fact that, in a later paper, Rutherford credits Bohr with drawing attention to this point (Rutherford 1914, p.498). On the other hand, in the 1911 article, Rutherford claimed: 'the question of the stability of the atom proposed need not be considered at this stage, for this will obviously depend upon the minute structure of the atom, and on the motion of the constituent charged parts' (Rutherford 1911, p. 671).

In proposing a model of a certain system, scientists do not commit themselves to the model being a completely faithful epistemic representation of the target system. It is only through an investigative process that our *competence* in using a certain model as a faithful epistemic representation of the system increases. This process consists in determining to what extent the valid inferences from the model to the system are sound. In some cases, some of the inferences that are valid according to the standard interpretation of the model may be found to be unsound when that aspect of the system is empirically investigated (as in the case of the Thomson model of the atom and Rutherford's scattering). In other cases, inferences whose conclusions are known to be incorrect may be shown to be validly inferred from the model (as it is in the case of Bohr's "discovery" of the instability of the atom in the Rutherford model).

I take these examples to show that scientific models are often used to perform inferences about a certain target system and, as such, they can be considered epistemic representations of their target systems. The Bohr model of the hydrogen atom and the 'H' letter on the periodic table both denote the atom, but only Bohr model of the atom is an epistemic representation of it. As I have argued, denotation is not sufficient for epistemic representation. This seems to suggest that, if there is no special problem about scientific representation, it is not because scientific representation is simply an instance of denotation as Callender and Cohen seem to think, but because scientific models are epistemic representations like the London Underground map.[10]

[10] Some may feel that there are still important differences between scientific models rand other epistemic representations that make the problem of scientific representation a special problem. (I will discuss a few of them in Section II.1.2.4 below.) From a methodological point of view, however, it seems reasonable to assume that scientific models represent their target exactly like any other epistemic representation until proven otherwise—until we can specify an essential difference in how models and other epistemic representations are used as epistemic representations.

II. Epistemic Representation

II.1. The Inferential Conception of Scientific Representation

II.1.1. INTRODUCTION

In this chapter, I will examine one of the main conceptions of scientific representation in the literature—Mauricio Suárez's inferential conception of scientific representation—and argue that it does not provide us with a satisfactory account of "scientific" representation. According to Suárez, the primary aim of a *substantial* conception of scientific representation—i.e. a conception which specifies non-trivial necessary and sufficient conditions for scientific representation—is to answer the question: 'By virtue of what is a certain vehicle a scientific representation of a certain target?' and not the question: 'By virtue of what is a certain vehicle an accurate or truthful scientific representation of that target?' In the terminology adopted here, the inferential conception of representation proposed by Suárez is, thus, meant to be a conception of "scientific" representation and not of what makes a "scientific" representation more or less faithful.

According to the inferential conception, '*A* represents *B* only if (i) the representational force of *A* points to *B* and (ii) *A* allows competent and informed agents to perform inferences regarding *B*' (Suárez 2004, p.773). Condition (i) seems to amount to what so far I have called denotation. Suárez maintains that '[...] this feature would be satisfied by a mere stipulation of a target for any source [which is Suárez's term for what I call 'vehicle']' (Suárez 2004, p.771). However, according to Suárez, in order for *A* to be a scientific representation of *B*, it is not sufficient that *A* denotes *B*: '[...] if a representation is to be objective in this sense (i.e., if it is to be a candidate for a scientific representation) it cannot be an arbitrary sign' (Suárez 2004, p.772). Hence, the second necessary condition for scientific representation: *A* is a scientific representation of *B* only if an informed and competent user would be able to perform (valid though not necessarily sound) surrogative inferences from *A* to *B*. Condition (ii) seems to amount to what I have called valid surrogative reasoning.

In the terminology that I have adopted here we can therefore reformulate the inferential conception as maintaining that

(A) A vehicle is a scientific representation of a target (for a user) only if:

(A.1) The user takes the vehicle to stand for the target and

(A.2) The user is able to perform valid surrogative inferences from the vehicle to the target.

Conditions (A.1)and (A.2) amount to what I have called, respectively, denotation and valid surrogative reasoning. According to the inferential conception, thus, denotation and valid surrogative reasoning are individually necessary but not jointly sufficient conditions for scientific representation.

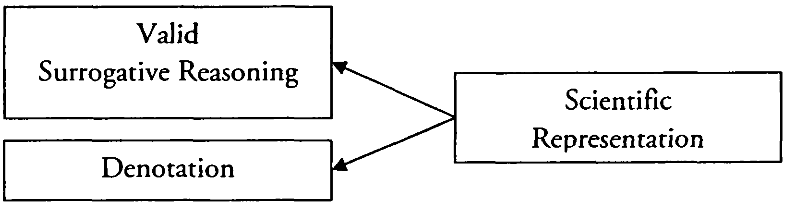


Diagram 2. The Inferential Conception

Suárez however does not think that this is a problem for the inferential conception of scientific representation because the inferential conception is not meant to be a substantial conception of scientific representation. According to him, 'representation is not the kind of notion that requires, or admits, [universal necessary and sufficient] conditions' (Suárez 2004, p.771). Thus, according to Suárez, a conception of scientific representation cannot provide us with a set of non-trivial conditions that are individually necessary and jointly sufficient for scientific representation.

Suárez offers two possible interpretations of the non-substantiality of the inferential conception. On the one hand, Suárez claims that one should not look for further conditions because there are '[...] no deeper features to scientific representation other than its surface features' (Suárez 2004, p.769). According to Suárez, these features are surface features in the sense that they are features of the *concept* of scientific representation (Suárez 2004, n. 4). On this interpretation, Suárez claims, the inferential conception would be a *deflationary* conception of scientific representation (Suárez 2004, pp.770–771).

On the other hand, Suárez seems to think that there are further, more concrete conditions by virtue of which the concept of scientific representation applies to cases of scientific representation, but that these further conditions differ from case to case. For example, Suárez claims that: 'in every specific context of inquiry, given a putative target and source, some stronger conditions 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, p.776).[11] According to this interpretation, Suárez would be claiming that a conception of representation can spell out a set of necessary conditions for the concept of scientific representation but it cannot spell out a set of necessary and sufficient conditions for its application. On this interpretation, the inferential conception would be, Suárez says, a *minimalist* conception of scientific representation.

In this chapter, I will mainly focus on the points on which I disagree with Suárez. Before turning to those, it is worth briefly mentioning a few of the most crucial points on which I agree with him. The first point on which I agree with Suarez is that the question 'By virtue of what does a certain model represent a certain target?' is to be clearly distinguished from the question 'By virtue of what does a certain model represent its target faithfully?'. The second point is that denotation is a necessary but not sufficient condition for epistemic representation. The third point of agreement is that valid surrogative reasoning is a necessary condition for epistemic representation.

[11] Similarly, see also (Suárez 2003, p.768 and p.776).

Despite my agreement with Suarez on these three points, however, there are a number of issues on which I disagree with him. The first point of disagreement concerns the relation between the conditions that I have called 'denotation' and valid surrogative reasoning. Suárez seems to think of them as two independent conditions, each of which is individually necessary for epistemic or scientific representation and neither of which implies the other. I think that Suárez is mistaken on this point—denotation and (possible valid) surrogative reasoning are not distinct conditions because denotation itself is a necessary condition for surrogative reasoning. For the moment I will put aside this point of disagreement because my reasons for thinking that denotation and surrogative reasoning are not independent will become clear only in the next chapter in which I develop and defend my account of epistemic representation.

The second point of disagreement concerns whether or not valid surrogative reasoning is a sufficient condition for epistemic or scientific representation. On the inferential conception, valid surrogative reasoning and denotation are necessary but not sufficient conditions for scientific representation. So, *a fortiori*, valid surrogative reasoning is not sufficient for scientific representation. On the definition of epistemic representation that I have proposed (definition (1) in Section I.2.1) valid surrogative reasoning is both necessary and sufficient for epistemic representation. So, an advocate of the inferential conception must deny that valid surrogative reasoning is sufficient for epistemic representation or that scientific representation is just a kind of epistemic representation. In Section II.1.2, I will argue that there seem to be no good reason to deny either of those claims.

The third point of disagreement concerns the relation between epistemic representation and the user's ability to perform valid surrogative inferences. In Section II.1.3, I will argue that, first, the fact that the user is able to perform surrogative inferences from the vehicle to the target is a consequence of the fact that the vehicle is an epistemic representation of the target (for that user), not the reverse and that, second, a satisfactory account of epistemic representation should explain by virtue of what a user is able to perform valid surrogative inferences from the vehicle to the target if the vehicle is an epistemic representation of the target. On the inferential conception, however, it would seem that the fact that the user is able to perform surrogative reasoning from the vehicle to the target is a brute fact that cannot be further explained or analysed. This, I think, makes the user's ability to perform valid inferences form the vehicle to the target unnecessarily mysterious.

II.1.2. Is Surrogative Reasoning Sufficient for "Scientific" Representation?

According to the inferential conception of scientific representation, surrogative reasoning is a necessary but not sufficient condition for scientific representation. This means that, according to the inferential conception, a user may able to perform inferences from a certain vehicle to a certain target without the vehicle being a scientific representation of that target. In this section I will consider a few reasons why one might think that this is the case and I will find them all wanting. My argument will be divided into two stages. First, I will argue that there seems to be no reason to think that surrogative reasoning is not sufficient for epistemic representation (II.1.2.1–II.1.2.3). Then, I will argue that there seems to be no reason to think that there are conditions that distinguish scientific representations from other epistemic representations (II.1.2.4).

II.1.2.1 Must Every Epistemic Representation Be Partially Faithful?

If surrogative reasoning was not sufficient for representation, then it would be possible for a user to perform surrogative inferences from the vehicle to the target even if the vehicle was not an epistemic representation of the target. Therefore, the best way to argue that (valid) surrogative reasoning is not sufficient for epistemic representation is to produce an example in which this is the case.[12] Suppose that someone performs valid inferences from the Rutherford model of the atom to a certain physical system, say, a hockey-puck sliding on the surface of a frozen pond.

[12] Again, although I have not defined what a valid inference from a model is, I shall take it to be sufficiently clear to proceed here, awaiting Section II.2.6 for further discussion of it.

From the model, for example, the user infers that the puck is negatively charged and the ice is positively charged, or that the puck's trajectory is circular. Since, on my characterization of epistemic representation, valid surrogative reasoning is sufficient for epistemic representation, this example seems to force me to accept the apparently problematic conclusion that the Rutherford model is an epistemic representation of the puck-on-ice system (or more precisely that it is one for the user in question).

Now, it is very likely that, from the model, the user will only draw 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 the user is able to perform valid inferences from the model to the system and not that any of these inferences are sound. Those who want to deny that (valid) surrogative reasoning is sufficient for epistemic representation, however, might want to claim that, if *all* the valid inferences from a certain vehicle to a certain target are unsound, then the vehicle is not an epistemic representation of the target. In other words, they might want to maintain that at least *some* of the valid inferences from the vehicle to the target need to be *sound* in order for the vehicle to be an epistemic representation of the target.[13]

If this were the case, however, a vehicle would be an epistemic representation of a certain target only if it was a partially faithful epistemic representation of the target. This, however, seems to be too strong a requirement, as it would rule out that completely unfaithful epistemic representations of a target are epistemic

[13] Whereas Suárez never embraces this position, Daniela Bailer-Jones (2003) seems to do so. If my interpretation of her brief remarks on the topic is correct, Bailer-Jones thinks that a certain model represents a certain system only if some of the conclusions about the system one can draw form the model are true. A great deal of (Bailer-Jones 2003) is devoted to investigating which of the conclusions that can be drawn from a model have to be true of a system in order for the model to represent the system. Bailer-Jones tentatively concludes that this depends on what the intended function of the model is and on which aspects of the system the model is intended to be about. Ultimately, then, it is the users of the model who determine which of the conclusions drawn from the model have to be true of the system in order for the model to represent the system.

representations of that target, which seems to be incoherent because an epistemic representation of a certain target, no matter how unfaithful, is still an epistemic representation of it.

To this, those who deny that valid surrogative reasoning is sufficient for epistemic representation are likely to object that the notion of epistemic representation is a technical notion and that, as such, there is no correct way of defining it. So there is nothing incoherent in defining it so that a vehicle is an epistemic representation of a certain target (for a certain user) if and only if that user is able to perform valid surrogative inferences from the vehicle to the target and some of the valid inferences from the vehicle to the target are sound.

Whereas there is nothing wrong with this definition in and of itself, this definition runs into trouble if one believes that epistemic representation should be distinguished from faithful epistemic representation, as Suárez and I do. Once we draw the distinction between epistemic representation and faithful epistemic representation, it becomes apparent that being an epistemic representation is a precondition for being a compeltely unfaithful epistemic representation of a certain target no less than it is a precondition for being a completely faithful epistemic representation of it. As I have already argued, when we fail to acknowledge this distinction, we can be misled into believing that we are contradicting ourselves when we say that something both represents and misrepresents something else.

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 that is intended to represent. Even if, gradually, we might come to discover that all the valid inferences from the model to the system are unsound, the model does not thereby cease to be regarded as an (epistemic) representation of the system. At most it ceases to be regarded as a *faithful* epistemic representation of the system. Once we distinguish the question of epistemic representation from that of faithful epistemic representation (as Suárez himself does), we come to realize that misrepresentation seems to presuppose representation no less than completely faithful epistemic representation. As I have already argued in Chapter I.2, a vehicle has to represent a certain target in order to *mis*represent it—or, in a less confusing terminology, a vehicle needs to be an epistemic representation of a certain target in order to be a completely unfaithful epistemic representation of it.

II.1.2.2Must Every Epistemic Representation Be Known to Be Partially Faithful?

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 to be all unsound. If this is right, then a vehicle represents a target only if its user does not know that all the inferences from the vehicle to the target are unsound. If, in the above example, the user knew that all inferences from the model to the system were 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 depends on the knowledge the user has of the system. However, in the above example, we have not made any mention of the knowledge that the user has of the system. As far as we know, the user could even mistakenly believe that some of the inferences from the model to the system 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. So, there would be nothing intrinsically wrong with claiming that, at least for that user, the Rutherford model is an epistemic representation of the puck-on-ice system.

Consider again the completely unfaithful model example that I have mentioned in the previous subsection. When the model was originally proposed, we did not know that none of the valid inferences from the model to the system were sound. Thus, for us, the model was an epistemic representation of the system. Then, we gradually discovered that all the inferences from the model to the system are, in fact, unsound. However, it seems absurd to suggest that we cease to regard the model as an epistemic representation of the system after we find out that all the inferences from the model to the system are unsound. Obviously, we cease to regard the model as a faithful epistemic representation of the system, but not as an epistemic representation of the system.

II.1.2.3Must Every Epistemic Representation Be Believed to Be Partially Faithful?

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 believes or has ever believed that they can draw sound inferences from the model to the system. This suggestion presupposes that a vehicle can be an epistemic representation of a target (for a certain user) only if at some point in time the user has believed that the vehicle is a partially faithful epistemic representation of the target, even if the vehicle is in fact a completely unfaithful epistemic representation of the system. The difference between this suggestion and the first is that, according to this suggestion, the vehicle may turn out to be a completely unfaithful epistemic 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 epistemic representation of the target but it might still be an epistemic representation of it if at some point the user believed it to be a partially faithful epistemic representation of the system.

Consider again the entirely unfaithful model example, at some point we might have believed that some of the inferences from the model to the system were sound, but this does not need to be the case. The model might have been proposed in a purely hypothetical and conjectural manner, without anyone necessarily believing that any of the inferences from it to the system were going to be sound. In other words, the model can be put forward purely as a generator of hypotheses about the system whose truth and falsity need 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-eighteen century.

So far, I have argued that there is no reason to deny that valid surrogative reasoning is sufficient for epistemic representation. If one were still to deny that valid surrogative reasoning is sufficient for epistemic representation, they would have to postulate that there is some "secret ingredient" that is present in the entirely unfaithful model case but is missing in the case of 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.

In fact, the difference between the two cases is, I think, circumstantial not substantial. According to the stories I have told, none of the conclusions about the puck-on-ice system validly drawn from the Rutherford model are interesting hypotheses about that system, while many of the conclusions validly drawn from the entirely unfaithful model were initially stimulating hypotheses about the system of which the model was an epistemic representation. The difference between the epistemic representations provided by the two models is not that one is an epistemic representation of its target system while the other is 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 completely unfaithful epistemic representations.

The difference is in the role the two models play in the context of the stories I have told about them. The investigation of the conclusions validly drawn from the completely unfaithful model may lead to new discoveries about the system and be instrumental to the development of more faithful models of that system, while the same would not apply to the case of the Rutherford model as an epistemic representation of the puck-on-ice system, if we already knew from the start that all of the valid inferences from the Rutherford model are unsound.

II.1.2.4 Can Scientific Representation be distinguished from Epistemic Representation?

So far I have argued that there seem to be no reason to deny that valid surrogative reasoning is sufficient for epistemic representation. However, what Suárez seems to deny is that valid surrogative reasoning is sufficient for *scientific* representation. If there were further conditions that distinguished scientific representations from other forms of epistemic representations, then surrogative reasoning may be sufficient for epistemic representation without being sufficient for scientific representation.

Before considering this option, let me note that it is very unlikely that this what Suárez actually thinks. Suárez repeatedly uses examples of epistemic representations that are not scientific in order to make points about "scientific" representation and does not seem to consider this practice problematic in any way. For example, one of Suárez's arguments against the similarity conception of representation is that the similarity between a portrait and a person is neither necessary nor a sufficient condition for the portrait to represent that person. If scientific representation was essentially different from other forms of epistemic representation, then the fact that similarity is not a necessary condition for other forms of epistemic representation would not show that it cannot be necessary for scientific representation (for all we know, even if Suárez's arguments are right with regards to the portrait case, similarity could still be a necessary condition for an epistemic representation to be a scientific one). In general, it would seem that, if one believes that cases of scientific representation were essentially different from cases of epistemic representation, they would engage in the practice of drawing lessons about scientific representation from non-scientific examples much more cautiously than Suárez does.

At any rate, even if we assume that Suárez does believe that scientific representation is essentially different from other forms of epistemic representation, it is at best unclear whether he has good reasons to think that there are any further conditions that distinguish scientific representations from other epistemic representations. Indeed it is not even clear which epistemic representations are scientific and which are not. For example, are geographic maps scientific representations or not? Many scientists, including geographers and geologists, use maps as part of their scientific investigations and the way they use these maps as epistemic representations does not seem essentially different from the way maps are ordinarily used as epistemic representations for non-scientific purposes. Analogously, electric circuit diagrams are used by physicists, engineers and hobbyists alike. Are the electric circuit diagrams that physicists use scientific representations and the ones that hobbyists use non-scientific epistemic representation?

Even if we were able to draw such distinctions, there seems to be no good reason to assume that the sense in which "scientific" representations represent their targets is different from the sense in which other epistemic representations represent their targets. The fact that we are not able to find sufficient conditions to demarcate scientific representations from other epistemic representations might, therefore, simply be due to the fact that there are no such further conditions—scientific representations are nothing but epistemic representations that scientists use in the pursuit of their research. In fact, I think that we should simply refrain from talking of "scientific" representation altogether unless we are able to identify any essential difference between "scientific" representation and other epistemic representation.

Some might object that further conditions that distinguish scientific representations from other epistemic representation must be there even if we are currently unable to identify them. Someone may be able to draw (valid) inferences about a certain swing from a photo of that swing but that does not make the photo a model of the swing. This objection however is based on a misunderstanding of what one should expect from an account of epistemic or "scientific" representation. The fact that a user is able to draw inferences from a map to the city of Venice makes of the map an epistemic representation of Venice (for that user) and the fact that a user is able to draw inferences from a view of Venice to the city of Venice makes of the painting an epistemic representation of Venice (for that user). However, what makes of the map a map and what makes of painting a painting is not the fact that users can perform inferences from them to the city of Venice. In other words, an account of epistemic representation can account for what makes of a certain map a map of Venice and what makes of a certain painting a view of Venice but should not be required to account for what makes them, respectively, a map and a painting in the first place. It may be the case that something can be a map, a painting or a model only if that something is an epistemic representation of something else (for some user), but an account of epistemic representation does not need to explain what makes of a certain representation a map, a model or a painting.

The arguments in this section obviously do not prove that the condition I have called valid surrogative reasoning is sufficient for epistemic representation or that there are no further conditions that distinguish "scientific" representations form other epistemic representations. However, I think the burden of the proof is now on those who want to deny either of those two claims.

II.1.3. On the Relation Between Epistemic Representation and Surrogative Reasoning

In the previous section, I have argued that there seems to be no reason to deny that (the possibility of) surrogative reasoning is a necessary and sufficient condition for epistemic representation or that scientific representations are nothing but epistemic representations. So, a substantial conception of epistemic representation seems viable. Since surrogative reasoning seems to be not only necessary but also sufficient for epistemic representation, then the inferential conception may be, in fact, providing us with necessary and sufficient conditions for epistemic representation. In fact, in a recent paper with Albert Solé, Suárez goes as far as suggesting that denotation and surrogative reasoning may well be both necessary and sufficient for the definition of the *concept* of scientific representation, but that this does not exclude that some further condition (such as similarity or isomorphism) are met in each concrete *application* of that concept. Even if this were the case, argue Suárez and Solé, the inferential conception would not be a substantial conception of scientific representation because, even if they were both necessary and sufficient, denotation and surrogative reasoning would only be surface features of scientific representation and a substantial conception of scientific representation is one that identifies *non-trivial* necessary and sufficient conditions for scientific representation.

In this section, I will argue that even if this was the case, the inferential conception would still not be a satisfactory conception of epistemic representation. I think that the situation is best illustrated by using medical terminology. Consider the following (fictional) medical example. Suppose that everyone comes in contact with the measles virus develops measles and everyone who develops measles has come into contact with the measles virus and that whoever has measles develops Koplik spots and everyone who develops Koplik spots has measels. In other words, having come into contact with the measles virus and having Koplik spots are each necessary and sufficient conditions for having measles. However, they are not on the same level from an explanatory point of view. Koplik spots are only *a symptom of* measles. *The cause of* measles is having been in contact with the virus. A satisfactory account of someone contracting measles is one that mentions the fact that they contracted the disease by coming into contact with the virus not one that mentions the appearance of Koplik spots.

Valid surrogative reasoning, I think, may well be a necessary and sufficient condition for epistemic representation, but it is only a symptom of epistemic representation. While the possibility of surrogative reasoning always accompanies epistemic representation, epistemic representation has conceptual precedence over actual surrogative reasoning. One is able to perform inferences from the London Underground map to the London Underground network *by virtue of* the fact that the map represents (i.e. is an epistemic representation of) the network for them, not the reverse—the map *does not* represent the network *by virtue of* the fact that one can use it to perform valid inferences about the network. In fact, we would not even attempt to use a piece of glossy paper with coloured 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 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.

The actual performance of surrogative inferences is just a "symptom" that allows us to tell apart cases of epistemic representation from cases of mere denotation. The actual performance of a surrogative inference from the model to the system reveals that the vehicle is being used as an epistemic representation of the target. In this sense the relation between representation and surrogative reasoning is analogous to that between measles and Koplick spots in the above example. Whereas the spots may be a necessary and sufficient condition for having measles, this does not mean that one has measles *because* they have Koplick spots.

Suárez would probably agree with the claim that valid surrogative reasoning is only a symptom of scientific representation. This I think is the most profound reason why he takes the inferential conception not to be a substantial conception of "scientific" representation—valid surrogative reasoning is just a surface feature of epistemic representation (one which *minimally* distinguishes it from mere denotation). The main point of departure between Suárez and me is therefore that, unlike Suárez, I believe that a satisfactory conception of epistemic representation *can* and *should* do more than listing a number of necessary symptoms or surface features of scientific representation.

The main reason why the inferential conception is unsatisfactory is that it does not explain by virtue of what a certain vehicle is an epistemic representation of a certain target (for certain users) and how, as such, it can be used by those users to perform (valid) surrogative inferences about the target. On the inferential conception, the user's ability to perform valid surrogative inferences from a vehicle to a target would seem to be a primitive feature that cannot be further explained. 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. Moreover, the inferential conception does not provide us with any principled way to distinguish between valid and non-valid surrogative inferences.

A satisfactory account of epistemic representation, I think, should address these questions—i.e. it should be able to explain:

- by virtue of what a vehicle is an epistemic representation of the target (for certain users),
- how the fact that the vehicle is an epistemic representation of the target for those users enables them to perform valid surrogative inferences from the vehicle to the target, and

- what makes some of these inferences valid and others not valid.

In the next chapter, I intend to offer an account that provides us with answers to these questions, which I will call the interpretational conception of epistemic representation.

II.2. The Interpretational Conception of Epistemic Representation

II.2.1. Introduction

So far, I have maintained that it is a necessary and sufficient condition for a vehicle to be an epistemic representation of a certain target (for a certain user) that the user is able to perform (valid) surrogative inferences from the vehicle to the target. I have called this necessary and sufficient condition for epistemic representation (valid) surrogative reasoning. I have also argued that a conception of epistemic representation which stops at this level is unsatisfactory. Surrogative reasoning, I have argued, is only a symptom of epistemic representation and a substantial conception of epistemic representation should do more than listing the symptoms of epistemic representation—it should identify those conditions by virtue of which the vehicle is an epistemic representation of the target.

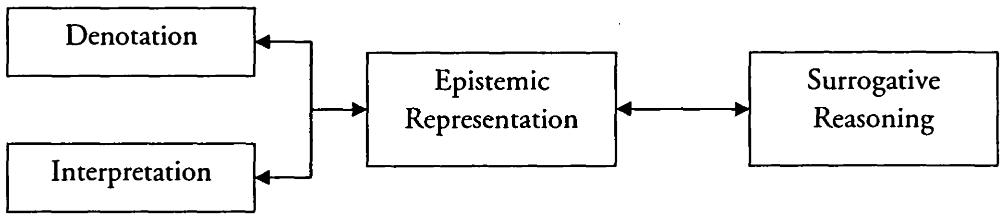
In this chapter, I will defend such a substantial conception of epistemic representation, which I call *the interpretational conception of epistemic representation.* The interpretational conception maintains that:

(B) A vehicle is an epistemic representation of a certain target (for a certain user) if and only if:

(B.1) the user takes the vehicle to denote the target, and

(B.2) the user adopts an interpretation of the vehicle.

Condition (B.1) is what, so far, I have called *denotation,* while condition (B.2) is what I will call *interpretation.*



According to the interpretational conception of epistemic representation, therefore interpretation is a necessary and sufficient condition for epistemic representation. The next two sections are devoted to clarifying what these two conditions amount to. I will first explain informally how when the two conditions that I have called denotation and interpretation hold, the vehicle is an epistemic representation of the target. In the following section, I will give a more formal definition of what it takes for a user to adopt a specific but very common kind of interpretation of a vehicle in terms of a target.

II.2.2. DENOTATION AND INTERPRETATION: AN EXAMPLE

Suppose that you are commissioned to design a map of a subway system, a map which represents which train lines connect which stations on the network. One way to go about this is the following. First, before even knowing what the network looks like, you identify what types of objects on the network and what types of properties of and relations among those types of objects are relevant based on what aspects of the subway system the map is supposed to represent. I will call these types of objects, properties and relations *T-relevant* (i.e. relevant in the target). Given that, in our example, the map is supposed to represent which stations are connected by the train lines, T-relevant types of objects must include the stations on the network, the T-relevant properties of those objects will include the names of those stations and the T-relevant relations among them will include the train services operating between those stations. However, not all objects, properties, functions, and relations are T-relevant. For example, if your map is not supposed to represent the relative positions of the stations, the positions of the stations will not be T-relevant properties of the type of object station.

The second step consists in selecting which types of objects on the map and which types of properties and relations among them will denote the T-relevant objects, properties and relations. For example, you can decide that, on your map, stations will be denoted by, say, small black circles, that a circle with a name printed on the side will denote the station with that name, and that direct train services operating between two stations will be denoted by coloured lines connecting the corresponding circles (each train line being denoted by a different colour). I will call *V-relevant* (relevant in the vehicle) the types of objects, properties, functions, and relations on the map that denote T-relevant types of objects, properties and relations.

After completing these two steps, you will have developed what I will call an *interpretation.* An interpretation has two interesting features. First, which types of objects properties and relations are T-relevant depends on what aspects of the target the epistemic representation is supposed to represent (depends on what I have called the scope of the representation). Were you interested in representing different aspects of the subway system you would have selected different types of objects, properties and relation from the one you actually chose as T-relevant. For example, if you were interested in designing a map of the subway system useful to the train drivers, the T-relevant objects would have probably included tracks, interchanges and platforms, rather than stations and train services.

Second, it is to a certain extent arbitrary what objects, properties, functions, and relations among objects on the map are V-relevant. Within the limits posed by pragmatic constraints, it would have been possible for you to use different types of objects, properties and relations on the map to denote the T-relevant types of objects, properties and relations. For example, nothing would have prevented you from using small red squares instead of small black circles to denote stations on the map. The pragmatic constraints, however, seem to set quite clear limits to the arbitrariness of your choices in this matters. For example, it seems to be highly impractical to use elephants to denote stations and extremely expensive to use precious stones. If we put aside the sheer impracticality (and animal rights), however, nothing would prevent you from producing an epistemic representation of a subway network in which stations are denoted by elephants and a direct trains service operating between two stations is denoted by the corresponding elephants being tied together by a coloured ribbon.

A third interesting feature of the interpretation discussed here is that it is, what I will call, a *general* interpretation—i.e. on the basis of it, one can construct a completely faithful epistemic representation of virtually any subway system. As I will argue later, not all interpretations have this feature.

Once you have developed a general interpretation, you can turn to designing the actual map of the specific subway system you are interested in based on the general interpretation. You will first turn to the actual subway system and make note of all and only those objects that are tokens of the T-relevant types of objects and of their properties and relations that are tokens of the T-relevant types of properties and relations. In our case, you compile a list of stations and make note of their names and of the train lines that connect them directly to other stations.

Then, you draw on your map one and only one small black circle for each station on your list and draw coloured lines between any two stations that are connected by a direct train service (using a different colour for each line). As a result of this process, you will have designed a map of the subway system in question. Here, I will call the interpretation on the basis of which you designed the map the *standard interpretation* of the map.

Consider now the map you have just designed from the perspective of one of its users. In and of itself, the map is not an epistemic representation of anything—it is just a piece of paper with small circles and coloured lines drawn on it. If the map is to become an epistemic representation for our user, the user must take some of the objects, properties and relations on that piece of paper to denote something else—i.e. the user has to adopt some interpretation or other of the map.

If the user is familiar with interpreting other maps of subway systems, it would probably be easy for her to realize that that piece of paper is meant to be a map of some subway system or other and that small black circles on the map denote the stations on that system and coloured lines between circles denotes direct train connections between the corresponding stations. In other words, a user who is familiar with similar interpretations of other subway maps is likely to work out what the standard interpretation of the map is.

However, the user does not need to adopt the standard interpretation of the map in order for that map to be an epistemic representation for her. For example, the user can take the circles to denote cities and towns, the coloured lines to denote highways, and the relative positions of the circles to denote the relative positions of the corresponding cities. Under such non-standard interpretations, the map would still be an epistemic representation for the user although it would not be an epistemic representation of a subway system but one of a highway system.

Assume that the user in question adopts the standard interpretation of your map. According to the interpretational conception, the map is now an epistemic representation of some subway system for that user—circles stand for stations and coloured lines stand for direct train connections. So it could seem that the condition I have called interpretation is not only necessary but also sufficient for epistemic representation. So, why does a second condition, denotation, enter the picture?

If the user only adopts an interpretation of the map without taking the map to denote any specific subway system, the map will only be a representation of some subway system or other not of any specific subway system. It is only when the user takes the map to stand for a specific subway system that the map becomes an epistemic representation of *that* subway system for that user. So for example, a user who adopts the standard interpretation of the map but does not take the map to stand for any specific subway system may infer that one of the stations on the subway network represented by the map is called Spadina, but she will not be able to infer that there is a station by that name on Toronto subway system unless they take that map to stand for Toronto subway network.

Somehow schematically, we could say that, when a user adopts an interpretation of a certain vehicle in terms of a certain type of target, the vehicle becomes an epistemic representation of a target of that type for that user, but it is only when the user takes that epistemic representation to stand for a specific target, that the epistemic representation becomes an epistemic representation of that target for that user. In our example, this piece of paper becomes a subway map for a user only when they interpret it in terms of some subway system or other; and the subway map becomes a *map of* Toronto subway system only when the user take the map (*qua* epistemic representation) to stand for the Toronto subway system.

II.2.3. ANALYTIC INTERPRETATIONS

According to a general, though somewhat loose, characterisation of the notion of interpretation, a user interprets a vehicle in terms of some target or other only if she takes facts about the vehicle to stand for (putative) facts about some target or other. One specific way to interpret a vehicle (though possibly not the only way) is to adopt what I will call an *analytic interpretation* of the vehicle. An analytic interpretation of a vehicle presupposes that the user identifies a (non-empty) set of V-relevant objects in the vehicle ($\Omega^V=\{o_1^V, \ldots, o_n^V\}$), a set of V-relevant properties of and relations among the V-relevant 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 V-relevant functions from $(\Omega^V)^n$—i.e. 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). The user assumes that there is a set of T-relevant objects in the target (Ω^T), a set of T-relevant properties and relations among objects in the target (P^T), and a set of T-relevant functions from $(\Omega^T)^n$ to Ω^T (Φ^T). (Here functions will be mainly used to stand for certain kinds of properties and relations such as *being 5.2 miles from*, which will be represented as a binary function that associates two objects with the real number 5.2, or *being 1.82m tall*, which will be represented as a unary function that associates an object with the real number 1.82)

(13) A user adopts an analytic interpretation of the vehicle in terms of the target if and only if:

[13.1] The user takes[14] 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,

[13.2] The user takes every *n*-ary relation in P^V to denote one and only one *n*-ary relation 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,

[13.3] The user takes 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.

What is peculiar about analytic interpretations is that relevant *objects* are denoted by *objects*, *properties* by *properties*, *binary relations* by *binary relations*, and so on. Note, however, that not all *prima facie* objects, properties, functions, and relations are relevant according to the interpretation of the vehicle in terms of the target. The only V-relevant objects, properties and relations are those that denote objects, properties and relations in the target and the only T-relevant objects, properties, functions and relations in the target are those that are denoted by, respectively, objects, properties, functions, and relations in the vehicle. For example, the relation *being connected by a light blue line* in the London Underground map is relevant (according to the standard interpretation of the map in terms of the network) because, on the standard interpretation of the map, it denotes a relation between stations on the network, but the relation *being two*

[14] Let me note that 'takes' here does not mean 'believes.' A user can take one object to denote another even if they do not believe there is anything the object denotes. I will say more about this below.

inches left of is not relevant because, on the standard interpretation of the map, it does not denote any relation among stations in the network.

II.2.4. ANALYTIC VS. NON-ANALYTIC INTERPRETATIONS

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 some objects on the map (i.e. small black circles and small coloured tabs) to denote objects on the network (i.e. stations). Third, we take some of the properties of and relations among those objects 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. In what follows, I will call any epistemic representation whose interpretation is analytic an *analytically interpreted epistemic representation.*

In this thesis, I will focus exclusively on what I have called analytic interpretations. Unless otherwise specified, by saying that a user adopts an interpretation of the target, I will always mean that the interpretation in question is analytic. However, I do not mean to imply that all interpretations of vehicles in terms of the target are analytic. Epistemic representations whose standard interpretations are not analytic are at least conceivable. For example, at a chess tournament with eight players, we may use a chessboard to keep track of which players are playing against each other at any given time. One letter from A to H and one number from 1 to 8 is assigned to each player so that every player is denoted by both a letter and a number and no two players are denoted by the same letter or the same number and if a piece on, say, the square D4 of the chessboard, denotes the fact that the player denoted by D is playing against the player denoted by 4. This is an example of an interpretation that is not analytic because a property of an object (i.e. the position of a piece on the chessboard) denotes a relation between two objects (i.e. the relation *playing against*).

Epistemic representations whose standard interpretations are non-analytic, however, 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.[15] If this is true, then, restricting our attention to analytic interpretations will extremely simplify the discussion without any comparable loss of generality.

Moreover, if one is willing to adopt a more liberal conception of what can count as relevant objects, properties and relations, it may be possible that any non-analytic interpretation can be reconstructed as an analytic one. In the example above, for example, we have considered a piece to be a V-relevant object and the position of a piece on the chessboard as a V-relevant property. However, if one was willing to consider columns and rows of squares on the chessboard to be V-relevant objects so that a chessboard is composed of 16 objects (8 rows and 8 columns of squares) and the relation *having the same piece on* a relation between a row and a column such that the relation holds between column *X* and row *Y* when and only when a piece is on one square *XY*, then the above interpretation becomes an analytic one because T-relevant objects (players) are denoted by V-relevant objects (rows and columns of squares) and a T-relevant binary relations (*playing against*) is denoted by a relation between rows and columns of squares.

The above is obviously not meant to be a general argument to the effect that any non-analytic interpretation can be reconstructed as an analytic one, but just to suggest that it is not implausible to think that such reconstruction may be possible in a wide variety of cases. At the moment, I do not have any general proof that any non-analytic interpretation can be reconstructed as an analytic one (in fact, I am not even sure what such a proof would look like). Whether the account I develop

---

[15] I talk of 'reconstruction' because, as I will mention later, 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.

in this thesis is a general account of epistemic representation and faithful epistemic representation crucially hinges on this question. If it was possible to reconstruct any interpretation as an analytic interpretation, then the account offered here would be a general account of epistemic representation. If this reconstruction was not always possible, I think that there are reasons to hope that, as it is, the account can deal with most prototypical cases of epistemic representation and that extending it to cases in which the interpretation adopted is irreducibly not analytic will only turn out to be technically challenging but not conceptually so.

II.2.5. Interpretation and Epistemic Representation

In ordinary language, we often use the same word to refer both to an epistemic representation and to the object that serves as a vehicle for that epistemic representation. For example, the word 'map' is sometimes used to refer to the material object that serves as a vehicle of a certain epistemic representation ('The map is on the table') and other times used to refer to the epistemic representation itself ('The map is very accurate'). However, the fact that we use the same word to refer to both should not mislead us into believing that the material object, in and of itself is an epistemic representation of something else. An epistemic representation is an object plus an interpretation of it.

According to the interpretational conception, it is only when a user, more or less consciously, adopts some interpretation or other of a vehicle that the vehicle becomes an epistemic representation for that user. This means a vehicle—by which I mean the object (material or not) that serves as a vehicle of a certain representation such as the piece of paper that serves as a map—in and of itself does not represent anything. It is only when some user adopts an interpretation of the vehicle that the object—the piece of paper—becomes an epistemic representation. According to this view, an epistemic representation can be schematically seen as an ordered pair $<V, I(V \rightarrow T)>$ whose first element is the vehicle (V, i.e. the object that serves as a vehicle of the epistemic representation) and whose second element is an interpretation of the vehicle in terms of the target ($I(V \rightarrow T)$).

Users are often unaware that they are adopting an interpretation of the vehicle in terms of the target. Even when they are aware of that, they would rarely be able to spell out exactly how they interpret the vehicle in terms of the target. For example, people are not usually aware that they adopt an interpretation to draw inferences from, say, a photograph to what the photograph represents. They feel that they can just *see* things *in* the photograph. This however does not mean that they do not in fact adopt an interpretation of the photograph but only that they are so used to interpreting photographs in a certain way that the interpretation process becomes transparent to them. It is only because we are so used to interpreting photographs that we can believe that we can directly "see" things in a photograph. It is only when we are less familiar with the standard interpretation of a certain form of epistemic representation (such as in the case of infrared photographs or ultrasound scans), that the need for an interpretation becomes apparent.

One of the reasons why interpretations often become transparent to us is that, in many cases, the same interpretation (or a family of closely related interpretations) can be used to interpret different vehicles as epistemic representations of different targets when these epistemic representations are representations of the same kind. So, we do not need to learn a new interpretation every time we come across a new epistemic representation if we are already familiar with the standard interpretation associated with the form of epistemic representation. For example, after learning how to interpret a geographic map, we are usually able to use the same (or a very closely related) interpretation for other geographic maps as well. I will call such interpretations that are used as the standard interpretation of many representations of the same form *general interpretations*.

This feature of interpretation finally puts us in the position to provide an answer to the question raised in Chapter I.1 concerning our preferences of the map of Michigan over the upturned right-hand and the upturned right-hand over the salt shaker as vehicles of an epistemic representation of the geography of Michigan. In the case of the map, we do not need to come up with an *ad hoc* interpretation of the vehicle in terms of the target because we are already familiar with a number of ready-made, general interpretations that allow us to interpret maps in terms of their target and these are likely to include one that will allow us to interpret the map of Michigan in terms of the state of Michigan (for example, we know that, usually, a blue area on the map denotes an expanse of water). In the case of the upturned right-hand, it is quite intuitive for those who are familiar with the geography of Michigan to come up with an interpretation of the upturned right-hand in terms of the geography of Michigan (one according to which a point on the palm of the hand denotes a certain location in Michigan and the contour of the hand denotes the borders and coasts of Michigan). However, there is no intuitive interpretation of the salt-shaker in terms of the geography of Michigan, let alone a ready-made one. According to the interpretational conception of epistemic representation, our preferences can thus be explained in terms of how easily we can interpret the different vehicles in terms of the targets.

II.2.6. How Does Interpretation Relate to 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. As we have seen, for example, 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. 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 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^{V}_{i} denotes o^{T}_{i} according to the interpretation adopted by the user, it is valid for the user to infer that o^{T}_{i} is in the target if and only if o^{V}_{i} is in the vehicle,

(Rule 2) If o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{V}_{n}, and ${}^{n}R^{V}_{k}$ denotes ${}^{n}R_{k}^{T}$ according to the interpretation adopted by the user, it is valid for the user to infer that the relation ${}^{n}R_{k}^{T}$ holds among o^{T}_{1}, …, o^{T}_{n} if and only if ${}^{n}R^{V}_{k}$ holds among o^{V}_{1}, …, o^{V}_{n},

(Rule 3) If, according to the interpretation adopted by the user, o^{V}_{i} denotes o^{T}_{i}, o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{V}_{n}, and ${}^{n}F^{V}_{k}$ denotes ${}^{n}F^{T}_{k}$, it is valid for the user to infer that the value of the function ${}^{n}F^{T}_{k}$ for the arguments o^{T}_{1}, …, o^{T}_{n} is o^{T}_{i} if and only if the value of the function ${}^{n}F^{V}_{k}$ is o^{V}_{i} for the arguments o^{V}_{1}, …, o^{V}_{n}.

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

(14) 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:

[14.1] it is in accordance with *(Rule 1)*, or

[14.2] it is in accordance with *(Rule 2)*, or

[14.3] it is in accordance with *(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) an analytic interpretation of a vehicle in terms of a target underlies a set of rules to draw valid surrogative inferences from the vehicle to the target (b) 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.

Before concluding this section, let me note that an analytic interpretation *underlies* a set of rules in the sense that the adoption of a certain set of rules is part and parcel of the adoption of the underlying analytic interpretation. So, for example, from the fact that there is a circle labelled 'Holborn' on the London Underground map one can validly infer that there is a station called Holborn if and only if one adopts an interpretation of the map according to which the map denotes the London Underground network and small circles with a name printed on the side denote interchange stations with that name.

II.2.7. Denotation and Epistemic Representation

The condition that so far I have called denotation is, therefore, a necessary condition for both epistemic representation and surrogative reasoning. In this section, I will clarify how my view of the relation between denotation and epistemic representation differs from the views of Suárez and Callender and Cohen on this issue.

Let us consider Suárez's view first. According to Suárez denotation and surrogative reasoning are both necessary conditions for epistemic representation. This would seem to suggest that Suárez considers these two conditions independent from each other. I think Suárez is wrong on this point. On the interpretive conception, denotation and surrogative reasoning are also necessary conditions for epistemic representation but they are not on the same level because they are not independent from each other—one, denotation, is a necessary condition for the other, surrogative reasoning. To see why the condition I have called surrogative reasoning cannot hold unless the condition I have called denotation holds, consider again the example of the user who takes this piece of paper to be a map of a subway system but has no idea of what subway system in particular the map is a map of. According to the interpretational conception, this is because, while the interpretation condition holds, the denotation condition does not hold—that is the user adopts an interpretation of the map in terms of some subway system or other but does not take it to be a representation of any target system in particular. If that user is to be able to perform inferences from the map to some specific target such as the London Underground network (i.e. if the condition that I have called valid surrogative reasoning is to hold), the user has to take the map to stand for the London Underground system and not, say, Toronto's subway system. So, the valid surrogative reasoning condition cannot hold unless the denotation condition also holds. The two conditions, therefore, are not independent as the denotation condition is a necessary (but not sufficient) condition for surrogative reasoning.

Consider Callender and Cohen's view of the relation between denotation and epistemic representation. According to the interpretational account, denotation is a necessary condition for epistemic representation and it also plays a crucial role in the condition that I have called interpretation. This, however, is a very far cry from claiming that epistemic representation *amounts to* denotation as the advocate of the denotational conception of epistemic representation, which I have discussed in Chapter I.1, do. The set of ink marks 'the Tube', for example, may denote the

London Underground network but, in and of itself, this does not make of them an epistemic representation of the network (for us). If 'the Tube' is not an epistemic representation of the London Underground network (for us), it is not because it fails to denote the London Underground network but because we do not adopt any interpretation of those marks in terms of the network—for example, we do not take the letter 'T' (or any other mark) to stand for an object on the London Underground network and we do not take the fact that the letter 'T' is taller than the letter 'u' to stand for a relation between the objects in the network denoted by those marks.

Let me note, however, that I am not denying that 'the Tube' *could be* an epistemic representation of the London Underground network. I am only denying that it *is* one. Nothing would prevent us from doing so, in which case, according to the interpretational conception of epistemic representation, the set of marks 'the London Underground network' would qualify as an epistemic representation of the network. In that case, however, 'the London Underground network' would be an epistemic representation of the network by virtue of the fact that we have adopted an interpretation of those marks in terms of the London Underground network.

Before concluding this section, let me note that a user does not need to believe that a certain object in the target exists in order to take an object in the vehicle to denote it according to the interpretation of the vehicle in terms of the target. For example, on the old London Underground map, there is a tab labelled 'Dover Street'. According to the standard interpretation of the map in terms of the network, that tab denotes a station whose name is 'Dover Street'. However, since there is no longer a Dover Street station on today's London Underground network, at present that tab on the old map fails to denote any station. Even if I know that that is the case, however, I can still take the tab labelled 'Dover Street' to denote the Dover Street station *according to the standard interpretation of the network* that I am adopting. In other words, if I adopt the standard interpretation of the map in terms of the network, I can take every circle and every tab on the map to denote a station *according to the standard interpretation of the map* even if I know that some circles and tabs fail to denote a station. (This is vaguely similar to the case in which I take the name 'Santa Claus' to denote Santa Claus according to a child who believes in the existence of Santa Claus even if I know that the name 'Santa Claus' fails to denote any real person).

A better way to put this is probably to distinguish between what *I* believe about the London Underground network and what *the map* (*under its standard interpretation*) "tells" me about the network—we can call this the *informational content of the map* (*under its standard interpretation*). If what I believe about the London Underground network conflicts with what the map "tells" me about the network, I will not conclude that my interpretation of the map is wrong, rather I will conclude that the map is an unfaithful epistemic representation of the network.

II.2.8. The Interpretational Conception and Scientific Models

According to the interpretational conception of epistemic representation, a model is an epistemic representation of a target system (for a user) if and only if the user adopts an interpretation of the model in terms of the system. We often go through the process of interpreting a model in terms of a certain system so effortlessly that we do not even realise that we are interpreting the model in terms of the system. For example, in the case of the inclined plane model and the soap-box derby which I have first considered in Section I.2.7.1, we take a simple model from classical mechanics and turn it into an epistemic representation of a certain system by interpreting it in terms of that system. As I believe it is usually the case with models, the interpretation is an analytic one—i.e. we take the inclined plane model to denote the soap-box derby system, we take the box in the model to denote one of the racers in the system, and we take some relevant properties of the box such as its position or its velocity to denote the position and velocity of the racer denoted by it.[16]

[16] Here I talk of the position of the box at a certain time and that of the racer as a property to keep in line with the language ordinarily used by philosophers. However, it is more convenient to

A more problematic case may seem that of models some of whose components fail to denote anything in the system represented by it. For example, according to the Ptolemaic model of the cosmos, the universe is a system of concentric crystal spheres. The Earth lies at the centre of the sublunary region, which is the innermost sphere. Outside the sublunary region are the heavens: eight tightly fit spherical shells, the outermost of which, the sphere of the fixed stars, hosts the stars. Each of the other spherical shells hosts one of the seven "planets," which, in the Ptolemaic model, include the Moon and the Sun. Each spherical shell rotates around its centre with uniform velocity. According to the standard interpretation of the model, the crystal spheres in the model denote the crystal spheres that supposedly host the planet in the system. However, one does not need to believe that there actually are crystal spheres that host the planets and stars in order to adopt an interpretation of the model according to which crystal spheres host the planets. Here it is useful to use again the distinction between the user's beliefs about the target and the informational content of a certain epistemic representation. The users of the model may believe that planets are not hosted by crystal spheres and nevertheless believe that what the model "says" is that the planets are hosted by crystal spheres.

To sum up, the interpretational conception of epistemic representation maintains that a vehicle is an epistemic representation of a certain target (for a certain user) by virtue of the fact (and if and only if) the user adopts an interpretation of the vehicle in terms of the target. An interpretation of the vehicle in terms of the target underlies a set of rules to draw inferences from the vehicle to the target. So, whenever a user adopts an interpretation of the vehicle in terms of the target they will be able to perform surrogative inferences from the vehicle to the target. At the same time, the fact that a user performs a (valid) surrogative inference

regard certain properties such as the position and the velocity of the box at a certain time as functions rather than properties. So the velocity function is a function that associates with a certain object its velocity at a certain time.

from the vehicle to the target presupposes that she adopts an interpretation of the vehicle in terms of the target, which underlies a set of rules to draw inferences from the vehicle to the target and therefore the fact that the user performs inferences from a certain vehicle to a certain target is a symptom of the fact that the vehicle is an epistemic representation of the target.

The adoption of an interpretation provides the user with a set of rules to perform inferences from the vehicle to the target. However, as I have argued, the fact that a certain inference is *valid*—i.e. that it is in accordance with the set of rules that the interpretation adopted by the user underlies—does not imply that the inference is *sound*—i.e. that its conclusion is true. Therefore, if the user adopts an interpretation of the vehicle in terms of the target, the vehicle is an epistemic representation of the target. However, this does not imply that the vehicle is a completely faithful epistemic representation of the target. Nor does it imply that it is a faithful epistemic representation at all. In fact, as I suggested in Section I.2.5, the vehicle may be an epistemic representation of the target and yet be a completely unfaithful epistemic representation of the target.

If the vehicle is to be to some degree a faithful epistemic representation of the target some further conditions need to hold. An account that identifies these further conditions is an account of faithful epistemic representation. Part III of this thesis is devoted to developing such an account.

III. Faithful Epistemic Representation

III.1. Similarity and Faithful Epistemic Representation

III.1.1. The Question of Faithful Epistemic Representation

In Chapter I.2, I have distinguished two main questions that are usually conflated:

1) By virtue of what is a vehicle an epistemic representation of a certain target?
2) By virtue of what is a vehicle a faithful epistemic representation of a certain target?

Both the inferential conception of scientific representation that I have discussed in Chapter II.1, and the interpretational conception of epistemic representation that I have developed in Chapter II.2 were meant to answer question (1). In the third part of this thesis, I will turn to question (2) and I intend to develop an account of what I have called faithful epistemic representation.

As I have already mentioned when introducing the notion of faithful epistemic representation in Section I.2.5, the faithfulness of an epistemic representation is a matter of degree. Overall an epistemic representation can be completely faithful, partially faithful or completely unfaithful. Moreover, an epistemic representation can be a more or less faithful epistemic representation of a certain target and, of two epistemic representations of a certain target, one may be more faithful than the other.

I take it that, in order to have a satisfactory answer to question 2), we need to have an answer to each of the following questions:

(i) In virtue of what is a completely faithful epistemic representation of a certain target completely faithful?

(ii) In virtue of what is a partially faithful epistemic representation of a certain target partially faithful?

(iii) In virtue of what is one of two epistemic representations of a certain target more faithful than the other (if it is)?

(iv) In virtue of what are two equally faithful epistemic representations of a certain target equally faithful?

In this third part of the thesis, I will formulate an account of faithful epistemic representation that not only answers the questions (i)–(iv), but also, in doing so, it reveals the systematic relations among these notions of epistemic representation.

Before turning to these questions let me note that there are a number of closely related questions that I will not address here. These questions include the following questions:

(v) In virtue of what is a specifically faithful epistemic representation of a certain target a specifically faithful epistemic representation of that target?

(vi) What justifies a user in believing that a certain epistemic representation is a faithful epistemic representation of that target?

(vii) What justifies a user in believing that a certain valid inference from the vehicle to the target is sound?

If I do not address any of these questions here, it is not because I do not think they are interesting or worthy question. To the contrary I think that they are very important questions. However I believe that questions (i)–(iv) have conceptual priority over questions (v)–(vii)—i.e. we need to have an answer to that set of questions in order to answer those in the second set of questions. Here, I will be solely concerned with what makes an epistemic representation of a certain target a faithful one, independently of what the user is justified to believe and of the circumstances in which the epistemic representation is used.

III.1.2. Reinterpreting the Similarity Conception and the Structural Conception as Conceptions of Faithful Epistemic Representation

Beside the inferential conception of scientific representation, the two main accounts of representation discussed in the literature on scientific representation are the similarity conception and the structural conception of representation. Unfortunately, the sympathizers of the similarity conception, which include Ronald Giere and Paul Teller, and those of the structural conception, which include Patrick Suppes, Bas van Fraassen, Steven French, and James Ladyman, have never formulated explicit and detailed accounts of scientific representation—i.e. accounts that provides us with a set of conditions for a vehicle to be a scientific (or epistemic) representation of its target or a faithful scientific (or epistemic) representation of it. This, however, should not particularly concern us here, where we are mainly interested in developing an account of faithful epistemic representation and not in determining what the above-mentioned authors actually say about scientific representation.

Somewhat ironically, the most clear formulation of the so-called *similarity* and *structural* conceptions of representation available in the literature is due not to someone who sympathizes with these views but to one of their critics—Mauricio Suárez (see in particular Suárez 2003). On Suárez's interpretation, the similarity conception and the structural conception would be conceptions of epistemic representation and, as such, they would be incompatible with the inferential and epistemic conception that I have considered in Part II. (Evidence of this is the fact that one of Suárez's main arguments in favour of the adoption of a deflationary or minimalist conception of representation is the failure of what he calls [iso] and [sim] in accounting for epistemic representation.).

According to Suárez's interpretation, the similarity conception of representation maintains that:

(C) A vehicle is an epistemic representation of a certain target if and only if:
(C.1) the vehicle is similar to the target

The structural conception maintains that:

(D) A vehicle represents a target if and only if:
(D.1) (The structure instantiated by) the vehicle is isomorphic to (the structure instantiated by) the target.

(C) and (D), I think, are misrepresentations of, respectively, the similarity conception and the structural conception of representation. I believe that, unlike the inferential conception, the similarity conception and the structural conception are conceptions of *faithful* epistemic representation rather than conceptions of epistemic representation—i.e. they aim at identifying which conditions need to hold for an epistemic representation of a certain target to be a *faithful* epistemic representation of that target.

On this interpretation, for example, the similarity conception, on which I will focus in this chapter, maintains that:

(E) A vehicle is a faithful epistemic representation of a certain target for a certain user if and only if:

(E.1) The vehicle is an epistemic representation of that target for that user and

(E.2) The vehicle is similar to the target in some respects and to some degree (where different versions of the similarity conception specify different respects and degrees of similarity).[17]

In saying this, I do not mean to claim that the sympathizers of the similarity conception and the structural conception *intend* or *mean* to account for faithful epistemic representation. In fact, they do not even distinguish between epistemic representation and faithful epistemic representation. Rather, I claim that, once we draw the distinction between epistemic representation and faithful epistemic representation, the best way to make sense of the similarity conception and the structural conception is to consider them as conceptions of faithful epistemic representation rather than epistemic representation *simpliciter*.

On this interpretation, the similarity conception and the structural conception, unlike the inferential and the interpretational conceptions, would be attempts to

[17] I refer the reader to III.2.1 for the general formulation of the structural conception of faithful epistemic representation.

answer question (2)—By virtue of what is an epistemic representation of a certain target an overall faithful one (insofar as it is)?—not question (1)—By virtue of what is a vehicle an epistemic representation of a certain target? As such, both of them presuppose that those who hold them adopt a conception of epistemic representation such as the interpretational conception that I have proposed in Chapter II.2 or the inferential conception proposed by Suárez. So, far from being incompatible with the inferential conception and the interpretational conception, they actually complement them.

In other words, the fact that a vehicle is an epistemic representation of a certain target is a necessary but not sufficient condition for it to be a faithful epistemic representation of that target. As I have argued in Section II.1.2, once we distinguish between epistemic representation and faithful epistemic representation, it becomes apparent that there may be epistemic representations that are completely unfaithful. As I interpret them, the similarity conception and the structural conception therefore attempt to identify which further conditions—i.e. which conditions beside the ones that make the vehicle an epistemic representation of the target—need to hold in order for the vehicle to be (to a certain degree) a *faithful* epistemic representation of the target.

On this interpretation of these two conceptions, it is neither necessary nor sufficient conditions for a vehicle to be an epistemic representation of a target that the vehicle is similar to the target or that the vehicle is isomorphic to the target. Rather, what the two conceptions claim is that, *if a certain vehicle is an epistemic representation of a certain target*, then it is a faithful epistemic representation of it in virtue of the fact that some further condition holds (such as the fact that the vehicle is similar to the target or the fact that a morphism holds between the structure of the vehicle and that of the target).

III.1.3. How the Standard Arguments Against the Similarity Conception Can Be Avoided

As I mentioned, my argument in favour of this interpretation of the similarity conception and the structural conception is not exegetical. I am not saying that this is the correct interpretation of what the sympathizers of the similarity conception or those of the structural conception actually say or think. As far as I am concerned, the sympathizers of the similarity conception and those of the structural conception could well have in mind what Suárez seems to think they have in mind. Rather, my argument in favour of the interpretation of structural conception and the similarity conception that I am proposing here is that, interpreted as conceptions of faithful epistemic representation rather than epistemic representation, the similarity conception and the structural conception avoid most of the objections that are usually levelled against them and indeed they can be turned into viable conceptions of epistemic representation. Here, I will only consider Suárez's arguments against (C) and argue that these arguments are ineffective when directed against (E). However, since the Suarez's arguments against (C) and (D) are analogous, the same considerations will apply *mutatis mutandis* to Suárez's arguments against (D) when directed against the structural conception of epistemic representation.

The first argument against (C), the logical argument, states that, while similarity is a reflexive and symmetric relation, representation is usually a non-reflexive, non-symmetric and non-transitive relation. A map is most similar to itself but it does not represent itself, it represents the city of which it is a map. If the map is similar to the city, then the city is similar to the map; however, the city does not represent the map. (E) does not claim that the map represents the city by virtue of their similarity but that, *if the map is a representation of the city*, it is a faithful one by virtue of the fact that the map and the city are similar (in certain respects and to a certain degree).

Indeed, according to (E), similarity is neither necessary nor sufficient for the map to be an epistemic representation of the city. The fact that the similarity relation is reflexive and symmetric and the representation relation is usually neither reflexive nor symmetric does not have any bearing on the plausibility of (E), because according to (E), there is no connection between similarity and epistemic representation.

The second argument, the argument form mistargeting, states that the similarity conception does not allow for mistargeting. By 'mistargeting,' Suárez means those cases in which a user mistakenly believes one object to be the target of a representation, while another object is its actual target. In Suárez's own example, a friend of his disguises himself as Pope Innocent X and this misleads Suárez into believing that Velasquez's portrait of Pope Innocent X represents his friend. However, despite the similarity between the portrait and the friend, the portrait does not represent the friend.

According to (E), however the similarity between Suárez's friend and the portrait of Pope Innocent X is neither necessary nor sufficient for the portrait to be an (epistemic) representation of Suárez's friend. Rather, (E) claims that, *if the portrait represented Suárez's friend*, it would be a faithful epistemic representation of him if and only if it was similar to him in certain respects and to a certain degree. However, since the portrait does not represent Suárez's friend, according to (E), no amount of similarity between the two in and of itself can turn the portrait into an (epistemic) representation of Suárez's friend.

The third and the fourth arguments, the non-necessity and non-sufficiency arguments, state that similarity is neither a necessary nor a sufficient condition for representation—i.e. something can be a representation of something else even if it is not similar to it and something can be similar to something without being a representation of it. However, (E) does not claim that similarity is either necessary or sufficient for epistemic representation. According to (E), similarity is a further condition that needs to obtain in addition to the conditions that make a vehicle an epistemic representation of the target if the vehicle is to be a *faithful* epistemic representation of it.

III.1.4. THE SIMILARITY CONCEPTION OF FAITHFUL EPISTEMIC REPRESENTATION

Once we are clearer about the interpretation of the similarity conception and the structural conception, however, we have to determine whether some version of them can be developed into a satisfactory account of faithful epistemic representation. This is what I will do in this and the next two chapters. In the remainder of this chapter, I will briefly consider the similarity conception of faithful epistemic representation and will argue that, even if there is no conclusive reason to reject the similarity conception as inadequate, such a conception does not provide us with a satisfactory framework for a *general* account of faithful epistemic representation. In the next chapter, I will then turn to the structural conception and develop what I take to be a satisfactory general account of faithful epistemic representation, which I will call the structural similarity account. The structural similarity account not only will avoid the problems that other structural accounts of faithful epistemic representation run into, but it also vindicates the main intuition, which underlies the similarity conception—i.e. that similarity (although a very abstract sort of similarity) is somehow involved in faithful epistemic representation.

In other words, I do not think that the similarity conception of faithful epistemic representation is mistaken. To the contrary, I believe that there is more than a grain of truth to the similarity conception of epistemic representation. If I favour a version of the structuralist conception of faithful epistemic representation, it is because I believe that not only does it provide us with a much more solid and well-defined general framework for understanding faithful epistemic representation but also because it allows us to understand how similarity is related to faithful epistemic representation to the extent that it is.

III.1.5. THE RELEVANT SIMILARITIES ACCOUNT

In the context of the literature on models and representation, the similarity conception has been mainly supported by Ronald Giere. According to Ronald

Giere, the relationship between scientific models and real-world systems is mediated by what Giere calls *theoretical hypotheses* (cf. Giere 1985, 1988). Theoretical hypotheses are part of the theory to which a certain model belongs and have the following general form:

> The designated real system is similar to the proposed model in specified respects and to a specified degree (Giere 1985, p.80).

As an example of a specific theoretical hypothesis, Giere mentions the following hypothesis:

> The position and velocities of the earth and moon in the earth-moon system are very close to those of a two-particle Newtonian model with an inverse square central force (Giere 1988, p.81).

In this specific theoretical hypothesis, the earth-moon system is a real world system, the two-particle Newtonian model is the model that represents that system, the respects of similarity are the positions and velocities of the relevant objects in the model and the systems and the degree of similarity is said to be very close.

According to the conception of representation that seems to be implicit in Giere's earlier works a certain model is a faithful epistemic representation of a certain system only if the theoretical hypothesis that mediates between the model and the system in question is true, or, in other words, only if the model is actually similar to the system in the respects and to the degree specified by the theoretical hypothesis. So, for example, the inclined plane model represents the soap-box derby faithfully only if there is a true theoretical hypothesis according to which the model is similar to the system in certain respects and to a certain degree.

But *in which respects* and *to what degree* does a model have to be similar to a certain system in order to represent the system faithfully? In what respects and to what degree, for example, has the inclined plane model to be similar to the soap-box derby in order for it to be a faithful epistemic representation of it? The advocates of what I will call the *relevant similarities account* seem to think that the answer to this question crucially depends on the purposes the users want to use the model for.

Giere, for example, claims:

> 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. Since scientists are *intentional* agents with goals and purposes, I propose explicitly to provide a space for purposes in my understanding of representational practices in science. So we are looking at a relationship with roughly the following form: *S* uses *X* to represent *W* for purposes *P*. (Giere 2004, p.743)

Paul Teller (2001, pp.401–402) claims:

> [...] once the relevant context has been specified, for example by saying what is to be explained or predicted and how much damage will result from what kinds of error, the needs of the case will provide the required basis for determining what kind of similarity is correctly demanded for the case at hand. More specifically, similarity involves both agreement and difference of properties, and only the needs of the case at hand will determine whether the agreement is sufficient and the differences tolerable in view of those needs. There can be no general account of similarity, but there is also no need for a general account because the details of any case will provide the information which will establish just what should count as relevant similarity in that case. There is no general problem of similarity, just many specific problems, and no general

> reason why any of the specific problems need be intractable. (Teller 2001, p.402)

The relevant similarities account maintains that:

(F) A vehicle is a specifically faithful epistemic representation of a certain system for a certain user (for certain purposes) if and only if:

(F.1) The model is an epistemic representation of that system for that user and

(F.2) The model and the system are similar in the relevant respects and to the relevant degree (where the relevant respects and degrees of similarity are determined by the purposes of the user).

A first thing to note about (F) is that, as I have formulated it, the relevant similarity account seems to be a general account of faithful epistemic representation. That is, it seems to be an account of what makes any epistemic representation (not only scientific models) into a faithful epistemic representation of a certain system. In fact, Giere does not seem to think that scientific models are a special case of representational devices and he often talks of epistemic representations that are not scientific models (such as maps or diagrams) in a way that suggests that they constitute a continuum with scientific models.

A second thing to note is that the relevant similarities account is primarily an account of *specifically* faithful epistemic representation. That is, it is trying to answer what makes a vehicle a faithful epistemic representation of the system for a certain specific purpose the user has. It is not clear whether the advocates of the relevant similarity account think that the notion of the overall faithfulness of a representation makes no sense or if they merely think that the notion of the overall faithfulness of a representation is a derivative notion.

I will now proceed to consider some of the alleged problems of the similarity conception. I will argue that whereas most of them may be overcome some of them seem to be inherent to the use of the notion of similarity.

III.1.6. The Vagueness of Similarity

Philosophers are usually wary of using the notion of similarity because similarity is an inherently vague notion. In the case of the similarity conception, this problem has been raised by many of its detractors who think that the appeal to similarity is trivial (because anything is similar to anything else) (e.g. Suárez 2003) or it is no more than a blank to be filled in on a case-by-case basis (e.g. Frigg 2002). One of the main tasks facing the advocates of the similarity conception is thus that of dispelling the impression that their use of the notion of similarity renders their account vacuous or uninformative.

The first step in performing this task is to note that the similarity conception of epistemic representation does not make use of the notion of overall similarity. According to the similarity conception (as formulated in (E)), the vehicle and the target do not need to be similar overall (whatever that means); they need to be similar *in certain specific respects* and *to a certain degree.* So, first of all, even if it is the case that virtually anything is similar to anything in some respect and to some degree, it is clearly not true that anything is similar to anything else in some respect and to some degree. Even if there is a sense in which my red notebook is similar to my silver laptop in some respects (say, they are both mid-sized material objects with a (roughly) flat parallelepipedal shape which usually are on my desk), they are clearly not very similar with respect to their colour (in fact one could say that they are dissimilar with respect to their colour).

More generally, whereas genuine overall similarity judgements[18] seem to be hopelessly vague, people seem to be fairly good at making restricted similarity judgements such as 'the two sweaters are very similar in colour' or 'the two players are similar in height' (by a 'restricted similarity judgement', I mean a judgement about the similarity of two objects to a certain degree in a certain context). This is not to deny that, in those contexts, the term 'similar' is to some degree vague and context-dependent, but that the vagueness and context-dependence that characterizes 'similar' is not more serious than the vagueness that characterizes terms such as 'tall.' Given a certain context, there are some clear cases of objects that are similar and objects that are dissimilar in certain respects and to a certain degree as, given a certain context, there are clear cases of people who are tall and people who are not.

III.1.7. The Link between Purposes and Respects of Similarity

The second step in dispelling the impression that the use of similarity renders the similarity conception vague or uninformative consists in formulating a general account of what respects and degrees of similarity are required given that the user wants to use the epistemic representation for a certain purpose. For example, according to the relevant similarities account, it would seem that in which respects and to which degree the inclined plane model needs to be similar to the soap-box derby system in order for it to be a specifically faithful epistemic representation of the system for the purpose of predicting the final velocity of the racer depends on

---

[18] Personally, I am inclined to doubt that we ever even try to perform genuine overall similarity judgments. I believe that when people seem to be talking about the overall similarity of two or more objects, they are actually saying that the objects are similar in the respects and to the degree implicitly specified by the context. If I tell you 'Don't buy that sweater—you already have a lot of similar sweaters', the respects and degrees of similarity are implicitly suggested by the context—that is, in this case, I probably mean that the sweaters in question are similar in colour, material or cut not that they are similar in being all on the planet Earth, being self-identical, or being not good conductors of electricity.

the purpose. In this case, for example, it would seem that the final velocity of the box should be similar to the velocity of the racer in order for the model to be a specifically faithful epistemic representation of that system. However, it is not clear whether this specific respect of similarity is sufficient to make the model a specifically faithful epistemic representation of the system. (It would seem that if all that matters is the similarity between the final velocity of the box and that of the car, then a model that says that the final velocity of the box is proportional to the length of the box is a specifically faithful epistemic representation of the soap-box derby if the final velocity of the racer happens to be similar to the final velocity of the box as determined by the model) Moreover, since what we want to ascertain is that the racer will not exceed a velocity we deem to be safe, we are more interested in the final velocity of the box not being lower than that of the racer rather than in it not being higher. This is because, for our purposes, a model that grossly underestimated the final velocity of the racers would be a less faithful epistemic representation than a model that overestimated it.

As this example seems to show, the relation between purposes and degrees and respects of similarity is a complex one. It is at best unclear whether any general account of how certain kinds of purposes (say prediction or explanation) relate to specific respects of similarity is possible. I am not claiming that such an account is impossible. What I am saying is that, in lack of it, however, the relevant similarity account can hardly be acquitted from the charge of vagueness. Its critics are likely to claim that the account remains hopelessly vague if what the relevant respects and degrees of similarity are is a question that can only be answered on a case-by-case basis.

III.1.8. Can Theoretical Models Be Similar to Concrete Systems?

Another fundamental problem for the similarity account of scientific representation is explaining in what sense a theoretical model, which is arguably not a concrete object, can be similar to a concrete system in the world. R.I.G. Hughes, for example, claims:

> [...] we may model an actual pendulum, a weight hanging by a cord, as an ideal pendulum. We may be even tempted to say that in both cases the relation between the pendulum's length and its periodic time is approximately the same, and that they are in that respect similar to each other. But *the ideal pendulum has no length, and there is no time in which it completes an oscillation* (Hughes 1999, p.S330; my emphasis).

I take it that, according to Hughes, only material objects—i.e. objects that are both actual and concrete—can have concrete properties such as having a certain length and completing an oscillation in a certain time; and, since, whatever the ideal pendulum is, it is not a material object, it cannot have a length and cannot oscillate. So, it would not be clear in what sense the ideal pendulum could be similar to the actual pendulum, which oscillates and has a certain length.[19]

The advocates of the similarity conception, however, do not need to maintain that the ideal pendulum *literally* has the concrete properties it is ordinarily said to have in order to claim that it is similar to a certain real-world pendulum. Consider what seems to be an analogous case. Whatever fictional entities, such as Sherlock Holmes, are, they are not concrete actual objects. Yet they are often said to have properties that only concrete objects have, such as smoking a pipe and living on Baker Street. Obviously, when people say things like 'Sherlock Holmes smokes a pipe' in a meta-fictional context (i.e. when talking about fiction), they do not mean that Sherlock Holmes literally smokes a pipe. Rather, they mean that "he" does so in some non-literal sense.

The fact that fictional characters do not literally have some of the properties they are ordinarily said to have does not seem to prevent us from comparing them to concrete actual objects. When discussing literary works, for example, people often compare fictional entities to actual concrete ones. Readers, for example, may discuss how closely the London of Dickens novels resembles the actual Victorian

[19] Analogous objections can also be found in (Suárez 2002) and (Callender and Cohen 2006).

London, or whether the historical Richard III was as ruthless as the homonymous character in the Shakespearian play, or how closely Sherlock Holmes resembles Dr. Joseph Bell, one of Conan Doyle's professors who allegedly inspired the Sherlock Holmes character. Readers can even note how a character in a novel reminds them of someone they know.

These similarity comparisons seem perfectly legitimate and do not seem to presuppose that the fictional entities whose characteristics are being compared with those of actual concrete ones are themselves actual concrete objects that literally have the concrete characteristics attributed to them. To object to those that compare the personality of Sherlock Holmes with that of Joseph Bell that they cannot do so because Sherlock Holmes is not an actual concrete person and as such does not have a personality would simply be to miss the point of that exercise. At no point do those who draw the comparison assume or need to assume that Sherlock Holmes *literally* has a personality in order to draw that comparison.

Let me note that here I do not mean to deny that accounting for how objects that are neither concrete nor actual can, in some sense, said to have concrete properties is a genuine philosophical problem. The ontological status of fictional entities is at the centre of a long philosophical debate. The available conceptions can be roughly grouped in three: those according to which there are no fictional entities, those according to which fictional entities are possible (concrete) objects and those according to which fictional characters are (actual) abstract objects. None of these conceptions of fictional entities is entirely uncontroversial but it is well beyond the scope of this section to discuss these options and their problems. Here it is sufficient to note that any empirically adequate account of what fictional entities are should be able to account for the fact that fictional entities are not actual concrete objects and yet are said to have concrete properties. In fact, this is all we need to account for the practice of assessing the similarities of fictional entities with real ones.

In saying this, I do not mean to belittle the task of explaining how an object that is not actual or concrete can have concrete properties in some sense. I only mean to show that, not only is the claim that non-concrete models are similar to concrete systems far from being self-refuting, but, in fact, it does not even raise a new philosophical problem to which there is no suggested solution. Theoretical models, such as the ideal pendulum or our inclined plane, can be therefore construed as fictional entities, which, in some sense, have the properties that are said to have, such as having a certain length and having a certain period of oscillation. As in the case of fictional entities, this seems to be sufficient to ground judgements of similarity between a theoretical model and a concrete system. To avoid the above problem, the advocate of the similarity account can, therefore, claim that theoretical models are fictional entities and adopt any of the available accounts of the ontology of fictional entities insofar as the account in question can account for the fact that fictional entities, in some sense, have the properties they are attributed.

The critics of the similarity account, however, may think that this is not sufficient to solve the problem. Whereas real-world pendulums have *definite* oscillation times and real strings have a *definite* length, the pendulum in the ideal pendulum has no definite oscillation time but only an *indefinite* oscillation time T and its string has no *definite* length but only an *indefinite* length L. So, how can the length of the string in a real-world pendulum be similar to the length of the string of an ideal pendulum, if the string of the ideal pendulum has no definite length?

Against this second objection, it should be noted, first of all, that, even if the ideal pendulum has no definite length and no definite oscillation time, it has *a* length and *an* oscillation time and the relation between them is a *definite* relation. Secondly, nothing prevents us from setting the values of any of the parameters of the ideal pendulum that have an indefinite value. This is what someone would do when they want to use the ideal pendulum model to represent a specific pendulum. She can, for example, set the parameter L so the string in the ideal pendulum so that its string is exactly 30 centimetres long. In the specified ideal pendulum, the string would thus have a *definite* length exactly like any real string. It is important to note that, usually, by setting some of the parameters of a model, one also thereby fixes other parameters. In the case of the ideal pendulum, for example, by setting the values of the length of the string, L, and the gravitational acceleration on the bob, g, one also indirectly fixes the period of the pendulum, which is equal to $2\pi(L/g)^{1/2}$.

III.1.9. The Similarity Conception of Faithful Epistemic Representation

One of the most deep-seated intuitions people ordinarily have about representation is that representation is closely related to similarity. For example, most non-philosophers would be happy to claim that a portrait is similar to its subject (in the sense that it resembles its subject) and that this similarity is somehow connected to the fact that the portrait represents its subject. Once this intuition is subjected to scrutiny, however, it becomes apparent that the relation is not as close as one might be tempted to think at first. Following Nelson Goodman's scathing criticism of this common intuition, most philosophers in the analytic tradition today are wary of drawing a direct connection between similarity and representation.

Once we distinguish between epistemic representation and faithful epistemic representation, however, one does not need to think that there is any relation whatsoever between similarity and epistemic representation in order for similarity to play an important role in representation. According to the similarity conception of faithful epistemic representation, similarity between the vehicle and the target (in certain respects and to certain degree) is a necessary but not sufficient condition for the vehicle to be a faithful epistemic representation of the target. So, for example, the similarity conception of faithful epistemic representation does not maintain that the portrait is an epistemic representation of its subject by virtue of their similarity. It claims that, if the portrait is an epistemic representation of the subject, it represents her faithfully if and only if the portrait and the subject are similar in specified respects and to a specified degree.

This conception of faithful epistemic representation may seem very tempting when applied to some prototypical cases of epistemic representation. One such case, for example, is that of (conventional) portraits. First, in that case, our intuitions about the similarity between the portrait and the subject seem to be reasonably clear. If we show someone a portrait of one of their friends, they will usually be able to tell us whether or not it "resembles" like their friend. Second, in those cases, the advocate of the similarity conception would seem able to tell a plausible story about why, if the portrait is a faithful epistemic representation of its subject, then its users will be able to perform sound inferences from the portrait to the subject. The story would go somewhat as follows. A portrait is a faithful epistemic representation of a certain only if it is similar to the subject (in certain respects and to a certain degree). For example, the portrait represents its target faithfully only if the colour of the patches of paint that stand for the irises of the subject are similar to a certain degree to the colour of the irises of the subject. Then, the user can infer (in accordance with the standard interpretation of conventional portraits) that the irises of the subject are of the same colour as the patches of paint that denote them. If the colour of the patch of paint is similar to the colour of the subject's eyes, the user's conclusion would be true (or at least approximately true).

The similarity conception, however, does not seem to do equally well in other cases. For example, it is far from obvious in which respects and to which degree, if any, the London Underground map is similar to the London Underground network. This is not to say that there is no sense in which they are similar. (In the next chapter I will specify in which sense they are similar insofar as they are). Rather it is to say that the two objects do not seem to be similar in the *intuitive* sense of the word. One is a piece of glossy paper with names, small circles, and coloured lines printed on it; the other is a system of trains, tracks, platforms, escalators, and so on. The two objects do not seem to be similar in any intuitive sense of the word. This seems to be a serious problem. As I have argued in III.1.6, the best strategy available to avoid the charge of vacuousness directed against the similarity conception is to appeal to our ordinary intuitions about the similarity of some objects in some respect and to some degree. However, in this case, the similarity between the two objects, if any, seems to be so abstract that our ordinary intuitions about similarity simply seem to fail us.

The advocate of the similarity account could claim that the map is similar to the network in that the circles and tabs on the map are connected by a coloured line and the stations denoted by those circles and tabs are connected by a direct train service, but this line of defence trades on the fact that we happen to use the same abstract verb ('connect') can be used to refer to two relations that do not seem to be similar at all. When we say that two stations are connected by a certain train line, we mean that certain trains travel back and forth between those stations. This is clearly not what we mean when we say that a line connects to points. In that case, nothing travels back and for the between those points.

Let me emphasize again that this is not to deny that the map and the network may be similar at some level. In fact, I do think that they are similar. Rather, I am claiming that, if the map and the network are similar, they are similar at a very abstract level—they are similar in a sense that, though somehow close to the intuitive sense of 'similar', needs to be carefully specified if the account is to clearly apply to cases such as this.

The notion of similarity, I think, does not afford us the technical resources to capture this abstract sense of similarity. In the next chapter, I will develop a version of the structuralist conception of representation that clearly specifies the abstract sense in which the map and the network are similar—in the terminology that I will introduce there the map and the network are structurally similar.

III.1.10. Conclusion

In this chapter, I have tried to defend the similarity conception of representation from some of the objections that are usually directed against it. As far as I can see, there are no conclusive objections against the similarity conception of faithful epistemic representation. There is no reason to believe that most of the difficulties considered here cannot be overcome from within the similarity conception. The last difficulty, on the other hand, can be avoided by claiming that the similarity conception is only meant to account for faithful scientific representation not for faithful epistemic representation *tout court*.

If I do not attempt to go down this route here, it is not because I believe that it is impossible to develop a defensible version of the similarity account that avoids completely the difficulties that I have discussed here. Rather it is because I think that it is possible to develop and defend an account of faithful epistemic that not only is more general but also vindicates some of the intuitions that seem to underlie the similarity account. In the next chapter, I will develop and defend this account.

As I have mentioned, I do not think that the similarity conception of faithful epistemic representation is mistaken. To the contrary, I believe that there is more than a grain of truth to the similarity conception of epistemic representation. If I favour a version of the structuralist conception of faithful epistemic representation, it is because I believe that not only it provides us with a much more solid and well-defined general framework for understanding faithful epistemic representation but also because it allows us to understand how similarity is related to faithful epistemic representation to the extent that it is.

III.2. Structure and Faithful Epistemic Representation

III.2.1. Introduction

In this chapter, I will defend a version of the structural conception of faithful epistemic representation that I will call the structural similarity account of faithful epistemic representation. Sympathizers of the structural conception of representation seemingly include the likes of Patrick Suppes, Bas van Fraassen, and Steven French. From the literature, one could gather the impression that the structural conception is a fully-formed view. Despite many insightful remarks and some systematic work, however, none of the sympathizers of the structural conception has developed a detailed account of epistemic representation or faithful epistemic representation—i.e. a set of conditions for a vehicle to be an epistemic representation or a faithful epistemic representation of a target—or provided us with any example of how this conception would apply to some concrete case.

In most cases, their views are confined to rather episodic and informal remarks. Even when they provide us with something closer to an account of faithful epistemic representation (as, for example, in (French and Ladyman 1999)), their views on representation remain largely unclear as they fail to provide us with worked-out example of how their account applies to some concrete instance of epistemic representation.

My most immediate goal in this chapter is therefore that of clarifying what the structural conception of representation could possibly be. Rather than trying to interpret what the sympathizers of the structural conception might mean when talking about representation, I will try to develop what I take to be the strongest version of the structuralist conception of faithful epistemic representation and illustrate how this applies to concrete cases of epistemic representation. My most ambitious goal is that of developing a successful version of the structuralist conception of epistemic representation.

The structural conception aims at accounting for epistemic representation in terms of a formal relation, a *morphism*, between set-theoretic structures. A *set-theoretic structure* is a n-tuple $S = <A^S, {}^mR_1^S, \ldots, {}^oR_j^S, {}^pF_1^S, \ldots, {}^rF_k^S>$, where A^S is a non-empty set of objects (which we will call the *universe* of S); ${}^mR_1^S, \ldots, {}^oR_j^S$ are *relations* on A^S (where the superscript before the R indicates the number of places of the relation and properties are construed as unary relations), ${}^pF_1^S, \ldots, {}^rF_k^S$ are *functions* from subsets of tuples of A^S (where the superscript before the F indicates the number of arguments of the function) to elements of A^S. (For the sake of simplicity, I will drop all subscripts and superscripts whenever they are not necessary)

Isomorphism and homomorphism are two common morphisms.

(15) A function, f, from the universe of **A** to the universe of **B** (f: $A^A \rightarrow A^B$) is a *homomorphism* if and only if:

[15.1] if $<a_1^A, \ldots, a_k^A> \in R_i^A$, then $<f(a_1^A), \ldots, f(a_k^A)> \in R_i^B$ for all R_i^A and R_i^B and

[15.2] $f(F_j^A(a_1^A, \ldots, a_p^A)) = F_j^B(f(a_1^A), \ldots, f(a_p^A))$ for all F_j^A and F_j^B.

(16) A function, f, from the universe of **A** to the universe of **B** (f: $A^A \rightarrow A^B$) is an *isomorphism* if and only if:

[16.1] For every $o_i^B \in A^B$, there is an $o_i^A \in A^A$ such that $f(o_i^A) = o_i^B$.

[16.2] For every $o_i^B \in A^B$, if $f(o_i^A) = o_i^B$ and $f(o_k^A) = o_i^B$, then $o_k^A = o_i^A$.

[16.3] For all R_i^A and R_i^B, $<o_1^A, \ldots, o_i^A> \in R_i^A$ if and only if $<f(o_1^A), \ldots, f(o_k^A)> \in R_i^B$.

[16.4] For all F_j^A and F_j^B, $f(F_j^A(a_1^A, \ldots, a_p^A)) = F_j^B(f(a_1^A), \ldots, f(a_p^A))$.

Two structures, **A** and **B**, are *homomorphic* if and only if there is a homomorphism from the universe of **A** to the universe of **B**; they are *isomorphic* if there is an isomorphism from the universe of **A** to the universe of **B**. Homomorphism and isomorphism are the two most common morphisms. However, there can be other morphisms. For the sake of simplicity, I will call two structures *x-morphic* if and only if they are homomorphic, or isomorphic, or some other morphism holds between them.

In Section III.1.2, I have argued that the structural conception should be considered a conception of *faithful* epistemic representation. On this interpretation, the different versions of the structural conception will have the following general form:

(G) A vehicle is a faithful (analytically interpreted) epistemic representation of a certain target for a certain user if and only if:

(G.1) the vehicle is an (analytically interpreted) epistemic representation of that target for that user and

(G.2) an *x*-morphism holds between the vehicle and the target.

It is important to note that, as it is, (G) is not an account of faithful epistemic representation but a general template from which specific versions of the structural conception of faithful epistemic representation can be generated. First, (G) does not specify what kind of faithful epistemic representation (G.1) and (G.2) are necessary and sufficient conditions for. (Is it completely faithful, partially faithful, or specifically faithful epistemic representation?). Second, according to (G.1), a vehicle is a faithful epistemic representation of a certain target for a certain user only if it is an epistemic representation of that target for that user. Therefore, any version of the structuralist conception of faithful epistemic representation needs to adopt an account of epistemic representation. For example, the advocates of the structuralist conception could adopt Suárez's inferential conception of epistemic representation. However, as it will become apparent, the interpretational conception of epistemic representation that I have developed is much better suited to the purposes of the structuralist conception than the inferential conception.

Third, (G.2) leaves unspecified two crucial details. The first is how a morphism which is defined as a relation between set-theoretic structures can hold between most vehicles and targets which, on the face of it, are not set-theoretic structures. The second is what specific morphism needs to hold between the vehicle and the target in order for the vehicle to be a faithful epistemic representation of the target.

In this chapter, I will develop what I take to be a successful version of the structural conception of faithful epistemic representation. In Section III.2.2, I will address the question of how vehicles and target that are not set-theoretic structures themselves can nonetheless be said to instantiate set-theoretic structures. This seems to be a crucial problem for the structural conception, which has been raised by some of its critics (see, for example, Frigg 2006) but has never been addressed by the supporters of the structuralist conception. In Section III.2.2, I will argue that the notion of interpretation provides us with a very natural way to reconstruct both the vehicle and the target as set-theoretic structures. I will call those structures the relevant structure of the vehicle and the relevant structure of the target (relative to a certain interpretation) and I will say that the vehicle and the target instantiate those structure (on a certain interpretation of them).

In Section III.2.3, I will address another crucial problem for the structuralist conception, which has never been adequately addressed. The relevant structure of the vehicle and that of the target may well be *x*-morphic, but many of the *x*-morphisms that hold between them are not sufficient for the vehicle to be a faithful epistemic representation of the target. To avoid this problem some supporters of the structuralist conception have claimed that the only *x*-morphisms that count are the intended ones (see, for example, van Fraassen 1997). Unfortunately, not much has been done in order to make the notion of an intended *x*-morphism anything more than an intuitive notion. In Section III.2.3, I suggest that, roughly, the intended *x*-morphims are those that associate an element of the relevant structure of the vehicle with one of the relevant structure of the target only if that element denotes the other element according to the interpretation of the vehicle in terms of the target adopted by the user.

In Section III.2.4, I will develop what I call the *intended isomorphism account* of completely faithful epistemic representation. The account is very useful to illustrate how the structuralist conception applies to the simplest cases of faithful epistemic representation—those in which the vehicle is a completely faithful epistemic representation of the target. Unfortunately, though, most cases of faithful epistemic representation are not cases of completely faithful epistemic representation. If the structuralist account is to be taken seriously, it has to account for partially faithful epistemic representation as well and, in those cases, an intended isomorphism fails to hold between the relevant structure of the vehicle and that of the target. Since our main concerns here are scientific models, this seems to be particularly crucial as I illustrate in detail in Section III.2.7. Given that most scientific models are to some extent idealized and approximated, they are rarely completely faithful epistemic representations of the system they represent. Nevertheless, most models are to some extent faithful epistemic representations of their target system. For example, even if geocentric models are largely unfaithful epistemic representations of the solar system, they are nonetheless often used as epistemic representations of it because many of the inferences that can be drawn from them about the apparent positions of the stars and the moon are sound. A conception of epistemic representation that was unable to account for epistemic representations that are only partially faithful would hardly apply to the case of scientific models.

This is a well-known problem among the supporters of the structural conception. The commonly-accepted solution to this problem is to opt for a morphism weaker than isomorphism as the morphism that holds between the relevant structure of the vehicle and the relevant structure of the target. Much of the disagreement concerns which morphism is adequate. In Section III.2.6, I distinguish three kinds of unfaithfulness—which I call incorrectness, incompleteness and inexactness—and I argue that any combination of them can characterize a partially faithful epistemic representation. In Sections III.2.9 and III.2.10, I consider some candidates for the role of the weaker morphism that holds between the relevant structure of the vehicle and that of the target and I argue that are inadequate as general accounts of partially faithful epistemic representation, for none of them can adequately account for epistemic representations that are both incorrect and incomplete.

In Section III.2.11, I then develop what I call the intended partial isomorphism account of partially faithful epistemic representation. This account is a refinement of the account proposed and defended by Steven French and his collaborators. There I show how the account, as developed here, is successful in dealing with epistemic representations that are both incorrect and incomplete. In Section III.2.11, I argue that the intended partial isomorphism account is also able to explain the success of the other accounts (including the intended isomorphism account) by showing how those morphisms are limit cases of intended partial isomorphism. In Section III.2.13, however, I will argue that the account is not equally successful in dealing with partially faithful epistemic representations that are inexact.

Finally, in Section III.2.14, I develop what I will call the structural similarity account of faithful epistemic representation and I will argue that this account can account for partially faithful epistemic representations that are incorrect, incomplete, and inexact. The notion of structural similarity which plays a central role in the account is introduced in III.2.15, where the faithfulness of an epistemic representation is explained in terms of the structural similarity of the relevant structure of the vehicle and that of the target (relative to a certain interpretation of the vehicle in terms of the target.)

III.2.2. Structural Reconstruction

One of the most fundamental problem that faces a structural conception of faithful epistemic representation is that morphisms are defined as relations between set-theoretic structures and, on the face of it, most vehicles and targets are *not* set-theoretic structures.[20] For example, neither the London Underground map nor the London Underground network is a set-theoretic structure. So, how would a conception of representation that relies on the formal notion of a morphism apply to objects such as the London Underground map and the London Underground network, which are not structures?

The most promising answer to this question, I think, is to maintain that, even if they are not set-theoretic structures, concrete objects or systems can *instantiate* set-theoretic structures. Two concrete objects or systems, then, could be said to be *x*-morphic if and only if they instantiate *x*-morphic structures. Roman Frigg (2006) has suggested that to say that a concrete object (or system of objects) instantiates a certain structure amounts to giving an abstract description of it that applies if and only if some suitable, more concrete description of it applies. Following Nancy Cartwright (1999, Ch. 2), Frigg suggests a necessary condition for one description to be more abstract than another. A description such as 'John is playing a game' is more abstract than a set of descriptions—such as one that includes 'John is playing chess', 'John is playing football', 'John is playing poker'—only if (a) 'John is playing a game' does not apply unless one or other of the concrete descriptions in that set applies *and* (b) the fact that the more concrete description is satisfied is what the fact that the more abstract description is satisfied consists in on that instance. Frigg then claims:

> [...] for it to be the case that *possessing a structure* applies to a system, *being an individual* must apply to some of its parts and *standing in a relation* to some of these. The crucial thing to realise at this point is that *being an individual* and *being in a relation* are abstract on the model of *playing a game.* (Frigg 2006, p.55)

---

[20] Some advocates of the semantic conception of theories seem to believe that scientific models *are* set-theoretic structures (see e.g. van Fraassen 1997). However, this would solve only half of the problem because the systems that models represent are not set-theoretic structures but concrete systems.

For example, the London Underground network instantiates a structure N=<{Acton Town, Aldgate, ..., Woodside Park, Woolwich Arsenal}, ...,$^{1}R_i^{N}${<Blackhorse Road>, <Brixton>, <Euston>, <Finsbury Park>, <Green Park>, <Highbury & Islington>, <Oxford Circus>, <Pimlico>, <Seven Sisters>, <Stockwell>, <Tottenham Hale>, <Vauxhall>, <Victoria>, <Walthamstow Central>, <Warren Street>},...> only if an abstract description that includes among other things 'there is a property that Blackhorse Road, Brixton, ..., Walthamstow Central, and Warren Street stations have and no other station has' is true of the London Underground network. I will call this description a *structural description* of the London Underground network. This structural description is true only if a suitable, more concrete description—such as the one that includes 'Blackhorse Road, Brixton, ..., Walthamstow Central, and Warren Street stations are on the Victoria Line and no other station is on the Victoria Line' or the one that includes 'Blackhorse Road, Brixton, ..., Walthamstow Central, and Warren Street stations have four escalators and no other station has four escalators'—is true of the network.

The structural description is more abstract because it is purely extensional and, as such, it does not tell us anything about the nature of the properties and relations among the objects in the universe of the structure, it only tells us that such and such objects in the universe of the structure have some property or other that no other object in the universe of the structure has or that such and such *n*-tuples of objects in the universe of the structure are in some relation or other and that no other n-tuple of objects in the universe of the structure is in that relation. In this example, the property could be that of *being on the Victoria Line*, but any other more concrete property (e.g.: *having four escalators*) would do if it is shared by the abovementioned stations and by no other stations on the network.

It is crucial to note that, on this account, the same object or system can instantiate a number of different structures depending on which more concrete description of the system we are basing the structural description on—i.e. depending on how we identify which of its parts are the objects in the universe of the structure and which of the concrete properties of and relations among those parts have their abstract counterpart included in the structure. The same vehicle and the same target thus are likely to instantiate not one but many structures.

This however is not a problem for the structural conception if one adopts the interpretational conception of epistemic representation that I have defended in Chapter II.2. There, I have argued that a vehicle is an epistemic representation of a certain target if and only if a user interprets the vehicle in terms of the target and that one way for a user to interpret a vehicle in terms of the target is to identify some objects, properties and relations in the vehicle that stand for objects, properties and relations in the target. In Section II.2.2, I have called these objects, properties and relations in the vehicle and the target, respectively, V-relevant and T-relevant (according to a certain analytic interpretation of the vehicle in terms of the target).

For example, I have claimed that the relation *being connected by a light blue line* in the map is V-relevant (according to the standard interpretation of the map in terms of the London Underground network) because, on the standard interpretation, that relation stands for a relation between stations on the network; the relation *being two inches left of*, on the other hand, is not V-relevant because, on the standard interpretation of the map, that does not stand for any relation among stations in the network. Analogously, the relation *being connected by Victoria Line trains* in the network is T-relevant because there is a relation in the map that stands for that relation, while the relation *being three miles away from* is not T-relevant because, on the standard interpretation, no relation on the map stands for that relation.

Now, suppose that there is a true description of the vehicle that includes all and only the objects and their properties and relations that are V-relevant (according to the interpretation of the vehicle in terms of the target) and that there is a true description of the target that includes all and only the objects and their properties and relations that are T-relevant (according to the interpretation of the vehicle in terms of the target). Under those ideal descriptions both the vehicle and the target instantiate a structure, which I will call, respectively, the *relevant structure of the vehicle*, V, and *the relevant structure of the target*, T, according to interpretation *I*.

In general, the notion of relevant structure of the vehicle and relevant structure of the target can be defined as follows:

(17) If V is a vehicle, T is a target, and $I(V \rightarrow T)$ is an analytic interpretation of the vehicle in terms of the target, V is the relevant structure of V (relative to I) if and only if:

[17.1] $A^V=\{o_1^V, \ldots, o_n^V\}$ is the universe of the structure V if and only if A^V is the set of V-relevant objects according to I.

[17.2] ${}^nR_k^V=\{<o_1^V, \ldots, o_n^V>, \ldots, <o_i^V, \ldots, o_j^V>\}$ if and only if some n-ary V-relevant relation holds among $o_1^V, \ldots, o_n^V$, among, ..., among $o_i^V, \ldots, o_j^V$, but does not hold among any other n-tuple of objects in A^V.

[17.3] ${}^nF_k^V\{f(o_1^V, \ldots, o_n^V)=o_k^V, \ldots, f(o_i^V, \ldots, o_j^V)=o_z^V\}$ if and only if some n-ary V-relevant function takes o_k^V as its value if its arguments are $o_1^V, \ldots, o_n^V$, ..., and takes o_z^V as its value if its arguments are $o_1^V, \ldots, o_n^V$.

(18) If V is a vehicle and T is a target and $I(V \rightarrow T)$ is an analytic interpretation of the vehicle in terms of the target, T is the relevant structure of T (relative to the interpretation I) if and only if:

[18.1] $A^T=\{o_1^T, \ldots, o_n^T\}$ is the universe of the structure T if and only if A^T is the set of T-relevant objects according to the interpretation I.

[18.2] ${}^nR_k^T=\{<o_1^T, \ldots, o_n^T>, \ldots, <o_i^T, \ldots, o_j^T>\}$ if and only if some n-ary T-relevant relation holds among $o_1^T, \ldots, o_n^T$, among, ..., among $o_i^T, \ldots, o_j^T$, but does not hold among any other n-tuple of objects in A^T.

[18.3] ${}^nF_k^T\{f(o_1^T, \ldots, o_n^T)=o_k^T, \ldots, f(o_i^T, \ldots, o_j^T)=o_z^T\}$ if and only if some n-ary T-relevant function takes o_k^T as its value for the arguments $o_1^T, \ldots, o_n^T, \ldots$, takes o_z^T as a value for the arguments $o_1^T, \ldots, o_n^T$.

So, on this account, an interpretation of the vehicle in terms of the target provides us with a principled and natural way to reconstruct the vehicle and the target as set-theoretic structures. An analytic interpretation of a vehicle in terms of the target singles out some objects, properties, functions, and relations in the vehicle and in the target as respectively V-relevant and T-relevant. It is natural to assume that, according to that interpretation, the relevant structure of the vehicle and of the target will therefore be the structures that include all and only those V-relevant and T-relevant objects and their *V*-relevant and T-relevant properties and relations.

The interpretational conception of epistemic representation and the notion of interpretation in particular turns out to be particularly well-suited to the purposes of the structural conception of epistemic representation as it provides a principled, natural way to turn a seeming weakness of the structural conception into a strength. Even when a vehicle and a target are not set-theoretic structure, if the user adopts an analytic interpretation of the vehicle in terms of the target, the vehicle and the target can be easily reconstructed as set-theoretic structures on the basis of that interpretation.

III.2.3. Intended Morphisms

In the previous section, I have introduced the notions of the relevant structure of the vehicle, V, and the relevant structure of the target, T, according to a given analytic interpretation. Now, a number of different morphisms may hold between the relevant structure of the vehicle and the relevant structure of the target. It is therefore crucial to identify a special class of morphisms, which I will call intended morphisms.

(19) A morphism, *f*, between the relevant structure of the vehicle and that of the target is *intended* (relative to a certain analytic interpretation of the vehicle in terms of the target *I*) if and only if:

[19.1] For all $o^{V}_{i} \in A^{V}$ and all $o^{T}_{i} \in A^{T}$, $f(o^{V}_{i})=o^{T}_{i}$ only if, according to I, o_{i}^{V} denotes o_{i}^{T};

[19.2] For all ${}^{n}R^{V}_{k}$ and ${}^{n}R_{k}^{T}$, $\langle o^{V}_{1}, \ldots, o^{V}_{n}\rangle \in {}^{n}R^{V}_{k}$, and $\langle f(o^{V}_{1}), \ldots, f(o^{V}_{n})\rangle \in {}^{n}R_{k}^{T}$ only if, according to I, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R_{k}^{T}$.

[19.3] For all ${}^{n}F^{V}_{k}$ and ${}^{n}F^{T}_{k}$, $f({}^{n}F^{V}_{k}(o_{1}^{V}, \ldots, o_{k}^{V}))={}^{n}F^{T}_{k}(f(o^{V}_{1}), \ldots, f(o^{V}_{k}))$ only if, according to I, ${}^{n}F^{V}_{k}$ denotes ${}^{n}F^{T}_{k}$.

According to the standard interpretation of the London Underground map in terms of the network, for example, a small circle or tab with a name printed on the side stands for the station with that name and the relation *being connected by a red line* stands for the relation *being connected by Central Line trains.* A morphism between the relevant structures of the map and the network is, therefore, intended only if it associates a circle or a tab with a certain name with the station with that name and it associates circles and tabs that stand in the relation *being connected by a red line* with stations that are in the relation *being connected by Central Line trains.*

It is important to note that not all morphisms are intended. Suppose, for example, that, by mistake, the printers have printed a map of the London Underground network that is completely identical to the regular map except for the fact that the names of the circle labelled 'Holborn' and the tab labelled 'Highbury and Islington' on the standard map have been inverted. If the standard map is isomorphic to the network, so is the defective map, for the regular map and the defective one are themselves isomorphic to each other. The isomorphism between the relevant structure of the defective map and that of the network is not intended though because it associates the tab labelled 'Holborn' with Highbury and Islington station and the circle labelled 'Highbury and Islington' with Holborn station, while, according to the standard interpretation of the map, a circle or a tab with a name printed on the side stands for the station with that name.

Consider another example. This time the printers have mistakenly inverted the colours of the red line and dark blue line on the map. If the regular map is isomorphic to the network, then the defective map will be isomorphic to the network as well. However, the isomorphism is not intended. According to the standard interpretation of the map, the relation *being connected by a red line* stands for the relation *being connected by Central Line trains* and the relation *being connected by a dark blue line* stands for the relation *being connected by Piccadilly Line trains.*

I think that no version of the structural conception of faithful epistemic representation can be successful unless it employs the notion of an intended morphism (or an analogous notion). If the morphisms between the relevant structure of the vehicle and that of the target is not intended, it may well be the case that all of the inferences from the vehicle to the target are wrong. To see why just think of a map of London Underground that is exactly like the actual map except for the fact that the colours of all the lines are randomly assigned and each tab or circle is randomly assigned the name of a station. Suppose that the defective map is such that the map is a completely unfaithful epistemic representation of the network—i.e. so none of the inferences from the map to the network that are valid according to the standard interpretation associated with the standard map is sound. (This seems to be possible if an appropriate combination of colours and labels is chosen). Since the standard map is isomorphic to the defective map, if the former is x-morphic to the network, the latter will also be x-morphic to the network.[21] However, whereas some of the x-morphisms that hold between the standard map and the network will be intended, none of the x-morphism between the defective map and the network will be. So, if we do not specify that the x-morphism between the relevant structure vehicle and that of the target is an intended morphism, two epistemic representations of the same objects may well be both x-morphic to the target without being both partially faithful epistemic representations of it.

Supporters of the structural conception have sometimes expressed the need to focus only on the intended morphisms between vehicle and target (see, for

[21] This follows from the fact that if A and B are isomorphic structures and A is x-morphic to a third structure C, then also B is x-morphic to C.

example, van Fraassen 1997). No account of what conditions a morphism needs to meet in order to be an intended one, however, is available in the literature. The notion of an analytic interpretation interpretation, however, allows us to clearly define what conditions a morphism needs to meet in order to be an intended one.

III.2.4. The Intended Isomorphism Account of Completely Faithful Epistemic Representation

In the previous two section, I have introduced two crucial notions—that of relevant structure and that of intended morphism. We now have the resources to formulate an account of completely faithful epistemic representation—the intended isomorphism account. This will mainly serve as an illustration of the kind of intuition that underlies the structural conception that a certain *x*-morphism needs to hold between the vehicle and the target if the vehicle is to be a faithful epistemic representation of the target.

According to the intended isomorphism account:

(H) A vehicle is a completely faithful (analytically interpreted) epistemic representation of a target for a user if and only if:
(H.1) the vehicle is an (analytically interpreted) epistemic representation of that target for that user and
(H.2) there is an intended isomorphism between the relevant structure of the vehicle and that of the target.

In order to show that the intended isomorphism account of completely faithful epistemic representation is successful (in all cases in which the epistemic representation is analytically interpreted), I need to show that:

(i) if the vehicle is an analytically interpreted epistemic representation of the target for a certain user and an intended isomorphism holds between the relevant structure of the vehicle and that of the target, then all valid surrogative inferences from the vehicle to the target are sound and
(ii) if the vehicle is an analytically interpreted epistemic representation of the vehicle in terms of the target and all valid surrogative inferences from the

vehicle to the target are sound, then an intended isomorphism holds between the relevant structure of the vehicle and that of the target.

Consider (i) first. If a user adopts an analytic interpretation of the vehicle in terms of the target, an inference from the vehicle to the target will be valid if and only if it is in accordance with the rules outlined in Section II.2.6. I will now show that, if an intended isomorphism holds between the relevant structure of the vehicle and that of the target, then all valid inferences—i.e. all inferences that are in accordance with *(Rule 1)*, *(Rule 2)*, or *(Rule 3)*—are sound.

(Rule 1). Assume that there is an object o_i^V in the vehicle and that o_i^V denotes o_i^T according to the interpretation adopted by the user. According to *(Rule 1)*, it is then valid to infer that there is an object o^T_i in the target. Now, if an isomorphism f holds between the relevant structure of the vehicle and the relevant structure of the target, then there must be an object, o_x^T, in the universe of the relevant structure of the target, A^T, such that $f(o_i^V)=o_x^T$. If f is an intended isomorphism however, it must be the case that $o_x^T=o^T_i$, because o^T_i is the object that is denoted by o^V_i according to the interpretation adopted by the user.

Now, assume that there is no object o^V_i in the vehicle and that o^V_i denotes o^T_i according to the interpretation adopted by the user. According to *(Rule 1)*, it is then valid to infer that there is no object o^T_i in the target. If an isomorphism f holds between the relevant structure of the vehicle and the relevant structure of the target then every object in the universe of the relevant structure of the target must be in one-to-one correspondence with some object different from o^V_i and, if this isomorphism is intended, then o^T_i cannot be among the objects in the universe of the structure of the target because an intended isomorphism would associate o^T_i only with o^V_i, which is the object that denotes it according to the interpretation adopted by the user. Therefore, it is sound to infer that the object o^T_i is not in the target. So, if an intended isomorphism holds between the relevant structure of the vehicle and that of the target, any inference that is in accordance with *(Rule 1)* will be a sound inference.

(Rule 2). Assume that certain V-relevant objects in the vehicle, o_1^V, …, o_n^V, are in a certain V-relevant n-ary relation, ${}^nR_k^V$ and that, according to the interpretation adopted by the user, o_1^V denotes o_1^T, …, o_n^V denotes o_n^V, and ${}^nR_k^V$ denotes ${}^nR_k^T$. According to *(Rule 2)*, it is therefore valid to infer that a relation ${}^nR_k^T$ holds among o_1^T, …, o_n^T. If an isomorphism f holds between the relevant structure of the vehicle and the relevant structure of the target, then a relation ${}^nR_x^T$ will hold among the objects $f(o^V_1)$, …, $f(o^V_n)$. If the isomorphism is intended, then, since o_1^V denotes o_1^T, …, and o_n^V denotes o_n^V, it must be the case that $f(o_1^V)=o_1^T$, …, $f(o_n^V)=o^T_n$ and, since ${}^nR_k^V$ denotes ${}^nR_k^T$, the relation ${}^nR^T_x$ must be the relation ${}^nR^V_k$.

Assume now that certain relevant objects in the vehicle, o^V_1, …, o^V_n, are not in a certain relevant n-ary relation, ${}^nR^V_k$ and that, according to the interpretation adopted by the user, o^V_1 denotes o^T_1, …, o^V_n denotes o^V_n, and ${}^nR^V_k$ denotes ${}^nR_k^T$. According to *(Rule 2)*, it is therefore sound to infer that the relation ${}^nR_k^T$ does not hold among o^T_1, …, o^T_n. Here there are two cases to consider. Either a different relevant n-ary relation holds between o^V_1, …, o^V_n or no n-ary relation holds among them. If a different n-ary relation holds among o^V_1, …, o^V_n and an isomorphism holds between the relevant structure of the vehicle and the relevant structure of the target, then there will be a relation ${}^nR_x^T$ that holds among the objects $f(o_1^V)$, …, $f(o_n^V)$. If the isomorphism is intended, then, since o_1^V denotes o_1^T, …, and o_n^V denotes o_n^V, it must be the case that $f(o_1^V)=o_1^T$, …, $f(o_n^V)=o^T_n$ but ${}^nR_x^T$ cannot be ${}^nR_k^V$ because, if the isomorphism is intended, $<o^V_1, …, o^V_n>\in R_k^V$, and $<f(o^V_1), …, f(o^V_n)>\in R_k^T$ only if, according to the interpretation adopted by the user, R_k^V denotes R_k^T. If no n-ary relation holds among o_1^V, …, o_n^V and an isomorphism holds between the relevant structure of the vehicle and the relevant structure of the target, then no relation holds among $f(o_1^V)$, …, $f(o_n^V)$. If the isomorphism is intended, , since o_1^V denotes o_1^T, …, and o_n^V denotes o_n^V, it must be the case that $f(o_1^V)=o_1^T$, …, $f(o_n^V)=o^T_n$ and therefore no relation holds among those objects. So, if an intended isomorphism holds between the relevant structure of the vehicle and that of the target, any inference that is in accordance with *(Rule 2)* is sound.

(Rule 3). Assume that the function ${}^{n}F^{V}_{k}$ has o^{V}_{i} as its value when its arguments are $o^{V}_{1}, \ldots, o^{V}_{n}$, and, according to the interpretation adopted by the user, o^{V}_{i} denotes o^{T}_{i}, o^{V}_{1} denotes $o^{T}_{1}, \ldots, o^{V}_{n}$ denotes o^{V}_{n}, and ${}^{n}F^{V}_{k}$ denotes ${}^{n}F^{T}_{k}$. According to *(Rule 3)*, it is therefore valid to infer that the value of the function ${}^{n}F^{T}_{k}$ for the arguments $o^{T}_{1}, \ldots, o^{T}_{n}$ is o^{T}_{i}. If an isomorphism f holds between V and T, then $f({}^{n}F^{V}_{k}(o^{V}_{1}, \ldots, o^{V}_{n}))=f(o^{V}_{i})={}^{n}F^{T}_{k}\ (f(o^{V}_{1}), \ldots, f(o^{V}_{n}))$. However, since, if the isomorphism is intended, then $f(o^{V}_{i})=o^{T}_{i}$ and ${}^{n}F^{T}_{k}\ (f(o^{V}_{1}), \ldots, f(o^{V}_{n}))={}^{n}F^{T}_{k}\ (o^{T}_{1}, \ldots, o^{T}_{n})$, it must be the case that ${}^{n}F^{T}_{k}\ (o^{T}_{1}, \ldots, o^{T}_{n})= o^{T}_{i}$. So, if an intended isomorphism holds between the relevant structure of the vehicle and that of the target, any inference that is in accordance with (Rule 3) is sound.

Since, if the user adopts an analytic interpretation of the vehicle in terms of the target, only inferences that are in accordance with *(Rule 1)*, *(Rule 2)*, or *(Rule 3)* are valid and all inferences that are in accordance with *(Rule 1)*, *(Rule 2)*, or *(Rule 3)* are sound if an intended isomorphism holds between the relevant structure of the vehicle and that of the target, we can conclude that, in all cases in which the user adopts an analytic interpretation of the vehicle in terms of the target, if an intended isomorphism holds between the relevant structure of the vehicle and that of the target, all valid inferences are sound.

Consider now (ii)—if the user adopts an analytic interpretation and all valid surrogative inferences from the vehicle to the target are sound, then an intended isomorphism holds between the relevant structure of the vehicle and that of the target. If a user adopts an analytic interpretation of the vehicle in terms of the target, then only surrogative inferences that are in accordance with *(Rule 1)*, *(Rule 2)*, or *(Rule 3)* are valid. I will therefore need to show that if all inferences that are in accordance with *(Rule 1)*, *(Rule 2)*, or *(Rule 3)* are sound then it must be the case that an intended isomorphism holds between the relevant structure of the vehicle and that of the target.

If all inferences that are in accordance with *(Rule 1)* are sound, then, it must be the case that, for every object o^{V}_{i} that is in the vehicle, the object denoted by o^{V}_{i}, o^{T}_{i}, is in the target and that, for every object o^V_i that is not in the vehicle, the object denoted by o^V_i, o^T_i, is not in the target.

If all inferences that are in accordance with *(Rule 2)* are sound, then it must be the case that, for every *n*-tuple of objects, $o^V_1, \ldots, o^V_n$, that are in a *n*-ary relation ${}^nR^V_k$, the objects denoted by $o^V_1, \ldots, o^V_n$, $o^T_1, \ldots, o^T_n$, are in the relation denoted by ${}^nR_k^T$ and that, for every *n*-tuple of objects, $o^V_1, \ldots, o^V_n$, that are not in a *n*-ary relation ${}^nR^V_k$, the objects denoted by $o^V_1, \ldots, o^V_n$, $o^T_1, \ldots, o^T_n$, are not in the relation denoted by ${}^nR_k^T$. So, an *n*-tuple of objects $o^V_1, \ldots, o^V_n$ is in a certain relation ${}^nR^V_k$ if and only if the objects denoted by $o^V_1, \ldots, o^V_n$ are in the relation ${}^nR_k^T$ denoted by ${}^nR^V_k$.

Finally, if all inferences in accordance with *(Rule 3)* are sound, then it must be the case that for every *n*-ary function ${}^nF^V_k$ whose value for the arguments $o^V_1, \ldots, o^V_n$ is o^V_i, the value of function denoted by ${}^nF^V_k$, ${}^nF^T_k$, is the object denoted by o^V_i when the arguments are the objects denoted by $o^V_1, \ldots, o^V_n$.

From this, it follows that, if all inferences in accordance with *(Rule 1)*, *(Rule 2)*, or *(Rule 3)* are sound, then it is possible to construct a function, *f*, from the relevant structure of the vehicle A^V to the relevant structure of the target of the target A^T such that:

a) For every o^V_i and o^T_i, $f(o^V_i)=o^T_i$ if and only if o^V_i denotes o^T_i according to the interpretation adopted by the user.
b) For all ${}^nR^V_k$ and ${}^nR_k^T$, $<o^V_1, \ldots, o^V_n> \in {}^nR^V_k$ if and only if $<f(o^V_1), \ldots, f(o^V_n)> \in {}^nR_k^T$ and ${}^nR^V_k$ denotes ${}^nR_k^T$ according to the interpretation adopted by the user.
c) For all ${}^nF^V_k$ and ${}^nF^T_k$, $f({}^nF^V_k(o^V_1, \ldots, o^V_n))={}^nF^T_k(f(o^V_1), \ldots, f(o^V_n))$ and ${}^nF^V_k$ denotes ${}^nF^T_k$ according to the interpretation adopted by the user.

Since a function that meets these conditions is an intended isomorphism, it follows that, if all inferences in accordance with *(Rule 1)*, *(Rule 2)* and *(Rule 3)* are sound, then an intended isomorphism holds between the relevant structure of the vehicle and that of the target.

III.2.5. An Example of Completely Faithful Epistemic Representation

I will now illustrate how the intended isomorphism account works with a concrete example. Suppose that an intended isomorphism holds between the relevant structure of the London Underground map and that of the London Underground network relative to the standard interpretation. This means that there is a bijective function that associates circles and tabs on the map with the stations they denote according to the standard interpretation of the map and that some circles or tabs have a certain property or are in a certain relation only if the stations they denote are in the property or relation denoted by that property. For example, if an intended isomorphism holds between the relevant structure of the London Underground map and that of the London Underground network, then the circle labelled 'Holborn' and the tab labelled 'Bethnal Green' are connected by a red line if and only if Holborn and Bethnal Green stations are connected by Central Line trains because, according to the standard interpretation of the map, the circle labelled 'Holborn' and the tab labelled 'Bethnal Green' denote respectively Holborn and Bethnal Green stations and the relation *being connected by a red line* denotes the relation *being connected by Central Line trains*. Therefore, it follows that, if the user were to infer from the map that Holborn and Bethnal Green stations are connected by Central Line trains in accordance with *(Rule 2)*, her inference would be sound and so for all other inferences that are valid according to the standard interpretation of the map.

It is important to note that all this holds only if the isomorphism between the relevant structure of the map and that of the network is an intended isomorphism. If no intended isomorphism held between the relevant structure of the map and that of the network, then some valid inferences would not be sound. Consider the examples discussed in Section III.2.3. The first is the map which is identical with the standard map except for the fact that on it the printers have mistakenly inverted the names of the circle labelled 'Holborn' and the tab 'Highbury and Islington'. If the standard map is isomorphic to the network, so is the defective map, for the regular map and the defective one are themselves isomorphic to each other. The defective map, however, is not a completely faithful epistemic representation of the network, as some of the inferences that one could perform from the map to the network are unsound. For example, from the defective map, it is valid to infer that Highbury and Islington station is on the Circle Line, when in fact Holborn is on that line but Highbury and Islington station is not. So, even if the defective map and the network are isomorphic, the defective map is not a completely faithful epistemic representation of the network because none of the isomorphisms between the relevant structure of the map and that of the network is an intended one.

Consider the case in which the printers have mistakenly inverted the colours of the red and dark blue line. The relevant structure of the defective map and that of the network are still isomorphic, but the isomorphism is not an intended one. As a consequence, the defective map is not a completely faithful epistemic representation of the network. According to the standard interpretation of the map, a user would validly infer that Piccadilly line trains operates between Holborn station and Liverpool Street station, while actually only Central Line trains do so. Despite the isomorphism between the relevant structure of the map and the network, the defective map is not a completely faithful epistemic representation of the network, as the isomorphism between the relevant structure of the defective map and the relevant structure of the network is not the intended one.

III.2.6. Partially Faithful Epistemic Representations and Three Kinds of Unfaithfulness

So far, I have argued that the intended isomorphism account of completely faithful epistemic representation is successful in all those cases of completely successful representation in which the user adopts an analytic interpretation of a vehicle in terms of a target, as in the case of the London Underground map. However, not all epistemic representations are completely faithful. In fact, scientific models, which are our main concern here, are usually far from being completely faithful epistemic representations of the systems they represent. In this section, I will distinguish three kinds of unfaithfulness that characterise partially faithful epistemic representations—incompleteness, incorrectness, and inexactness—and argue that the intended isomorphism conception cannot account for partially faithful epistemic representation.

(20) A vehicle *V* is an *incorrect* (analytically interpreted) epistemic representation of a certain target *T* for a certain user if and only if:

[20.1] *V* is an analytically interpreted epistemic representation of *T* for that user, and

[20.2] V is the relevant structure of *V* (relative to the analytic interpretation *I* adopted by the user), and

[20.3] T is the relevant structure of *T* (relative to the analytic interpretation *I*(*V*→*T*) adopted by the user), and

[20.4] :

[20.4.1] for some $o_i^{V} \in A^{V}$, o_i^{V} denotes o_i^{T} and o^{T}_i does not exist, or

[20.4.2] for some ${}^{n}R_k^{V}$, o^{V}_1 denotes o^{T}_1, …, o^{V}_n denotes o^{T}_n, ${}^{n}R^{V}_k$ denotes ${}^{n}R_k^{T}$ and $<o^{V}_1, \ldots, o^{V}_k> \in {}^{n}R^{V}_k$ but $<o^{T}_1, \ldots, o_n^{T}> \notin {}^{n}R_k^{T}$.

(21) A vehicle *V* is an *incomplete* (analytically interpreted) epistemic representation of a certain target *T* for a certain user if and only if:

[21.1] *V* is an analytically interpreted epistemic representation of *T* for that user, and

[21.2] V is the relevant structure of *V* (relative to the analytic interpretation *I* adopted by the user), and

[21.3] T is the relevant structure of *T* (relative to the analytic interpretation *I*(*V*→*T*) adopted by the user), and

[21.4] :

[21.4.1] for some $o^{\mathrm{T}}{}_{i} \in \mathrm{A}^{\mathrm{T}}$, $o^{\mathrm{V}}{}_{i}$ denotes $o^{\mathrm{T}}{}_{i}$ and $o^{\mathrm{V}}{}_{i}$ does not exist

or,

[21.4.2] for some ${}^{n}R_{k}{}^{\mathrm{T}}$, $o^{\mathrm{V}}{}_{1}$ denotes $o^{\mathrm{T}}{}_{1}$, …, $o^{\mathrm{V}}{}_{n}$ denotes $o^{\mathrm{T}}{}_{n}$, ${}^{n}R^{\mathrm{V}}{}_{k}$ denotes ${}^{n}R_{k}{}^{\mathrm{T}}$ according to I, and $<o^{\mathrm{T}}{}_{1}, \ldots, o_{k}{}^{\mathrm{T}}> \in {}^{n}R_{k}{}^{\mathrm{T}}$ but $<o^{\mathrm{V}}{}_{1}, \ldots, o_{n}{}^{\mathrm{V}}> \notin {}^{n}R_{k}{}^{\mathrm{V}}$.

(22) A vehicle *V* is an *inexact* (analytically interpreted) epistemic representation of a certain target *T* for a certain user if and only if:

[22.1] *V* is an analytically interpreted epistemic representation of *T* for that user, and

[22.2] **V** is the relevant structure of *V* (relative to the analytic interpretation *I* adopted by the user), and

[22.3] **T** is the relevant structure of *T* (relative to the analytic interpretation $I(V \rightarrow T)$ adopted by the user), and

[22.4] For some for some ${}^{n}F^{\mathrm{V}}{}_{k}$, $o^{\mathrm{V}}{}_{1}$ denotes $o^{\mathrm{T}}{}_{1}$, …, $o^{\mathrm{V}}{}_{n}$ denotes $o^{\mathrm{T}}{}_{n}$, ${}^{n}F^{\mathrm{V}}{}_{k}$ denotes ${}^{n}F^{\mathrm{T}}{}_{k}$ and ${}^{n}F^{\mathrm{V}}{}_{k}(o^{\mathrm{V}}{}_{1}, \ldots, o^{\mathrm{V}}{}_{k}) \neq {}^{n}F_{k}{}^{\mathrm{T}}(o^{\mathrm{T}}{}_{1}, \ldots, o_{n}{}^{\mathrm{T}})$.

(Incidentally, it is worth noting how each kind of unfaithfulness is defined in terms of the failure of one of one of the conditions that need to hold in order for an intended isomorphism to hold between the relevant structure of the vehicle and that of the target).

In Chapter I.2, I have illustrated the difference between completely faithful and partially faithful epistemic representation by means of two maps of the London Underground network—a contemporary map of the London Underground and an old, 1930s map of it. As I noted, whereas, under its standard interpretation, the new London Underground map is a completely faithful epistemic representation of today's network, the old London Underground map is only a partially faithful epistemic representation of it under its standard interpretation, as only some of the inferences from the map to today's network that are valid according to that interpretation are sound. I will now use that example to illustrate the difference between the first two kinds of unfaithfulness—incorrectness and incompleteness. (I will illustrate the third kind of unfaithfulness, inexactness, in Section III.2.7 below).

According to the intended isomorphism conception of completely successful representation, the old London Underground map is no longer a completely faithful epistemic representation of the London Underground network because no intended isomorphism holds between the relevant structure of the map and that of today's network. There are at least two ways in which the intended isomorphism between the relevant structure of the old map and that of the network fails to hold and as a consequence the old map is both an incorrect and an incomplete epistemic representation of today's network. The old map is an incomplete epistemic representation of today's network because, among other things, some of the stations and train lines on today's network have no counterpart on the map. For example, on today's network, Victoria Line trains operate between Highbury and Islington and Victoria station. In the relevant structure instantiated by today's network, the ordered pair <Highbury and Islington, Victoria> is an element of the set of all pair of stations connected by Victoria Line trains. However, the circles on the old map that denote those stations (according to the standard interpretation of the old map) are not connected by any coloured line. Therefore, the ordered pair <tab labelled 'Highbury and Islington', circle labelled 'Victoria'> is not an element of any set of pairs of circles or tabs connected by a coloured line. As a consequence, an intended isomorphism fails to hold between the structure of the old map and the structure of the network.

The old map is an *incorrect* epistemic representation of today's network because among other things, some of the circles, tabs, and coloured lines on the map have no counterpart in today's network. For example, on the old map there is a tab labelled 'Dover Street' connected by a dark blue line to the circles labelled 'Piccadilly Circus' and 'Green Park'. In the structure instantiated by the old map, the tab labelled 'Dover Street' is an element of the set of circles connected by a blue line. Since there is no station called Dover Street on today's network, however, the tab labelled 'Dover Street' fails to denote any station on today's network. Therefore, there is no intended isomorphism between the map and the network as an intended isomorphism is one that associates every circle on the map with the station with the name printed by the circle.

Despite the old map being an incomplete and incorrect epistemic representation, a large number of inferences from the map to today's network that are valid according to the standard interpretation of the map in terms of the network are sound. Users would therefore be able to draw a number of sound inferences from the map to the network without adopting a non-standard interpretation of the map in terms of the network. This is far from being an exceptional case. For example, the pre-modern users of the Ptolemaic model (with epicycles and deferents), who were unaware that the model (on its standard interpretation) was an incorrect epistemic representation of its target, were able to perform a number of sound inferences from it such as those concerning the apparent position of the Sun, the Moon, the planets and the stars in the sky. I take it that a satisfactory account of faithful epistemic representation should be able to explain by virtue of what partially faithful epistemic representations, such as the old map and the Ptolemaic model of the universe, are faithful insofar as they are.

An alternative approach would be to assume that, for every epistemic representation that is partially faithful under its standard interpretation, there is some *ad hoc* interpretation of the vehicle in terms of the target under which the vehicle is a completely faithful epistemic representation of the target. I take this alternative to be unsatisfactory for a number of reasons. The main reason is that this alleged *ad hoc* interpretation is not the one users actually adopt. For example, someone who is unaware that the old London Underground map is obsolete from it to the target adopting their standard interpretation and not an *ad hoc* interpretation such that all inferences that are valid according to that interpretation are sound. If we are to be able to explain why those users of those epistemic representations can draw *both* true *and* false conclusions about the target from them we have to explain by virtue of what the map is a faithful epistemic representation of the network *on its standard interpretation.*

III.2.7. AN EXAMPLE OF INEXACT EPISTEMIC REPRESENTATION

It is widely acknowledged that scientific models are usually far from being completely faithful epistemic representations of their target systems. Consider again, for example, the inclined plane and the soap-box derby. In the example, the users intend to use the model to estimate the velocity of the racers at the foot of the hill to determine whether the racers will exceed the velocity that they deem safe. Suppose that, once we have plugged in acceptable approximations of the values of the height of the start line and of the gravitational acceleration in the model, the final velocity of the box turns out to be lower than the velocity we deem to be safe. By virtue of what is this conclusion true, if true? The structural conception of faithful epistemic representation would maintain that it is true in virtue of the fact that a certain morphism holds between the structure instantiated by the model and that instantiated by the system under the description of them that underlies the interpretation of the model in terms of the target. In this section, I will argue that this intended morphism, however, cannot be isomorphism.

This is not a new or surprising result. It is well-known that, since most (if not all) scientific models are idealized and approximated epistemic representations of their target systems, isomorphism cannot be the morphism that holds between the relevant structure of a model and that of its target system. In fact, to my knowledge, not one of the sympathizers of the structuralist account thinks that isomorphism can be the morphism that holds between idealized models and the systems they are used to represent. However, I think it is instructive to see exactly how isomorphism fails to obtain, because it will help us to identify the characteristics that a morphism should have in order to play this role and it will illustrate the third kind of unfaithfulness mentioned above—inexactness.

III.2.7.1 The Relevant Structure of the Inclined Plane Model

Following the work of Patrick Suppes and his collaborators (see, e.g., Suppes 2002) and of Wolfgang Balzer, Ulises Moulines and Joseph Sneed (Balzer, Moulines and Sneed 1987) on the set-theoretic structure of the models of classical particle mechanics, it is plausible to maintain that, on its standard interpretation, the inclined plane model instantiates a structure of the general form $M=<<A^M, T^M, V, R, I^+> r^M, m^M, f^M, g^M>$, where A^M is a non-empty set of objects, which in our case contains the box, T^M is an interval of real numbers, V is a three-dimensional vector space over real numbers, R is the set of real numbers, I^+ is the set of positive integers. The domain of r^M is the Cartesian product of A^M and T^M, $A^M \times T^M$, and its co-domain is V. For every $a_i \in A^M$ and every $t_j \in T^M$, r^M is twice differentiable at t_j. For the sake of clarity, I will call $v^M(a_i, t_j)$ and $a^M(a_i, t_j)$ the first and the second derivative of $r^M(a_i, t_j)$ with respect to t ($dr^M(a_i, t_j)/dt= v^M(a_i, t_j)$ and $d^2r^M(a_i, t_j)/dt^2=a^M(a_i, t_j)$). Informally, the functions r^M, v^M and a^M associate each object and a time-value with, respectively, the magnitude of the position, velocity and acceleration of that object at that time. The domain of the function m^M is A^M and its co-domain is the set of positive real numbers, R^+. Informally, m associates every object in A^M with the magnitude of its mass. The domain of f^M is the product of $(A^M \times A^M) \times T^M \times I^+$ and its co-domain is V. Informally, f^M associates ordered pairs of objects in the domain, a time-value and a positive integer with one of the forces that the first object exerts on the second one at that time (the positive integer is just a way to label different forces that one object may exert on the other). These forces can be construed as "internal" forces—forces exerted by objects that are within the system in question). The domain of g^M is $A^M \times T^M \times I^+$ and its co-domain is V. The function g^M associates an object, a time-value and an integer with one of the "external" forces acting on it at that time (where external forces can be informally construed of as forces that are exerted by objects outside of the system in question or by source-less force fields).

There are different sort of constraints on the values these functions can have given a certain set of arguments. The most general set of constraints, which I will call the *general constraints*, stem from the fact that the inclined plane model is a model of classical mechanics and, as such, the objects in it are subject to the general laws of classical mechanics such as the Newtonian laws of motions, which constrain the values of the functions. For example, according to Newton's Second Law, for every $a_i \in A^M$ and every $t_j \in T^M$, $f^M(a_i, a_1, t)+\ldots+f^M(a_i, a_{i-1}, t)+f^M(a_i, a_{i-1}, t)+\ldots+f^M(a_i, a_{i-1}, t)+\sum_{k=1}^{\infty} g^M(a_i, t, k)=m^M(a_i)a^M(a_i, t)$. According to Newton's third law, for every couple of objects, a_i and $a_k \in A^M$ and every $t_j \in T^M$, in P, $f^M(a_i, a_k, t_j)=-f^M(a_k, a_i, t_j)$.

A more specific set of constraints, which I will call *specific constraints*, derives from the specific features of the inclined plane model. In this particular case, the specific constraints concern the values of the function g^M. In the model, two forces act on the box at all times. The first one is an external gravitational force, $g^M(b, t_i, 1)$, whose magnitude is constant and equal to $m^M(b)g$ (where g is the gravitational acceleration). The second is the normal force $g^M(b, t_i, 2)$ that the plane exerts on the box whose direction is perpendicular to the plane and whose magnitude is constant and equal to $m^M(b)g \sin\theta$ (where θ is the angle of inclination of the plane).[22] Since there are no other forces acting on the box, $g^M(b, t_i, k)=0$ for all $k>2$.

The general and specific constraints are all the constraints on the values of the functions in the structure of the inclined plane model as such. Despite these constraints, however, the functions in the structure of the model still do not have definite values for all arguments unless some additional constraints are put on the model. I will call this further set of constraints the *inputs of the model*. In the case of the inclined plane model, one such set of constraints consists in specifying the mass

---

[22] Note that here I consider the box the only object in the universe of the structure of the inclined plane model A^M and the normal force of the plane on the box an external force. This is because considering the plane an object itself would give rise to certain unintuitive consequences. For example, if the plane was one of the objects in the universe of the structure of the model, the function m^M would associate a certain mass with it and there is no obvious sense in which the plane in the model has a mass or a position.

of the box ($m^M(b)$), its initial position and velocity ($\mathbf{r}^M(b, t_0)$ and $\mathbf{v}^M(b, t_0)$), the gravitational acceleration (the values of g), and the angle of inclination of the plane (the value of θ). By specifying g and θ, the function $\mathbf{g}^M$ will have a definite value for all arguments. Once we specify ($m^M(b)$), $\mathbf{a}^M(b, t_i)$ will have a definite value for all arguments as well $\mathbf{a}^M(b, t_i)=(\mathbf{g}^M(b, t_i, 1)+\mathbf{g}^M(b, t_i, 2))/m^M(b)$. Now we need only to specify the position and velocity of the box at some t_i such as t_0 in order for $\mathbf{r}^M(b, t_0)$ to have definite values at all other times as well (($\mathbf{v}^M(b, t_i)=\mathbf{a}^M(b, t_i)t_i+\mathbf{v}^M(b, t_0)$ and $\mathbf{r}^M(b, t_0)=\frac{1}{2}\mathbf{a}^M(b, t_i)t_i^2+\mathbf{v}^M(b, t_0)t_i+\mathbf{r}^M(b, t_0)$). Until the inputs of the model are specified, the inclined plane model, therefore, does not instantiate a structure. Rather, it instantiates what following van Fraassen (, p.44) I will call a *structure-type.* I will call *structure-tokens* those structures that are instances of a certain structure-type. So, the inclined plane model instantiates a structure-token only when a specific set of inputs of the model are specified.

III.2.7.2 The Relevant Structure of the Soap-Box Derby System

Consider now the soap-box derby system. The system can be seen as instantiating a structure-type as well. Each token of that structure type is a structure of the form $S=\langle\langle A^S, T^S, V, R, I^+\rangle, \mathbf{r}^S, m^S, \mathbf{f}^S, \mathbf{g}^S\rangle$, where A^S is a non-empty set of objects that contans one and only one of the racers, T^S is an interval of real numbers, V is a three-dimensional vector space over real numbers, R is the set of real numbers, I^+ is the set of positive integers. The domain of $\mathbf{r}^S$ is the Cartesian product of A^S and T^S, $A^S \times T^S$, and its co-domain is V. The functions $\mathbf{r}^S$, $\mathbf{v}^S$ and $\mathbf{a}^S$ associate the racer in A^S and a time-value $t_i \in T^S$ with, respectively, the magnitude of its position, velocity and acceleration at that time on a certain run of the derby. The domain of the function m^S is A^S and its co-domain is the set of positive real numbers, R^+. The function m^S associates the racer in A^S with the magnitude of its mass. The domain of $\mathbf{f}^S$ is the product of $(A^S \times A^S) \times T \times I^+$ and its co-domain is V. Informally, $\mathbf{f}^S$ associates ordered pairs of objects in the domain, a time and an integer with one of the forces that the first exerts on the second one at that time. The domain of $\mathbf{g}^S$ is $A^S \times T \times I^+$ and its co-domain is V. The function $\mathbf{g}^S$ associates the racer in A^S and a time that has magnitude $t_i \in T^S$ with one of the external forces acting on it at that time on a specific run of the derby.

We can therefore think of the value of the functions in each token of the structure-type of the soap-box derby system as representing the value of that quantity at a certain time in a specific run of the system for a specific racer. This means that in each token of the structure-type of the system the function $\mathbf{r}^S(f(b), f(t_i))$ associates the j-th racer that takes part to the derby with its exact position at the time denoted by t_i on its k-th run ($\mathbf{r}^S(f(b), f(t_i))=\mathbf{r}^S(f(b), f(t_i))_{jk}$). I will call this token of the structure-type S, S_{jk}.

Since morphisms are only defined for structure-tokens, not structure-types, if we want to apply the notion of a morphism to a case like that of the inclined plane model and the soap-box derby, in which the relevant structures of both the vehicle and the target are structure-types, we have to make sense of the notion of a morphism between structure-types. Here, I will call the *inputs of the system* those aspects of the system that correspond to the aspects of the model that I have called the inputs of the model. For example, since the inputs of the inclined plane model include the mass, initial position, and initial velocity of the box, the inputs of the soap-box derby system will include the mass, initial position and velocity of the racer denoted by the box. On each run of each racer, these inputs will have definite values. Therefore, for each structure-token S_{jk}, the functions, $m^S(f(b))_{jk}$, the initial position, $\mathbf{r}^S(f(b), f(t_0))_{jk}$, the initial velocity, $\mathbf{v}^S(f(b), f(t_0))_{jk}$ etc. will have certain specific values. The token of the structure-type of the inclined plane model that corresponds to the structure-token S_{jk}, M_{jk}, is the one in which the value of all the inputs of the model is set equal to the value of the inputs of the system in structure-token S_{jk} (e.g. $f(m^M(b))=m^S(f(b))_{jk}$, $f(\mathbf{r}^M(b, t_0))=\mathbf{r}^S(f(b), f(t_0))_{jk}$, $\mathbf{v}^M(b, t_0)=\mathbf{v}^M(f(b), f(t_0))_{jk}$). The structure-type of the inclined plane model, M, and that of the soap-box derby system are thus x-morphic if and only if, for all M_{jk} and S_{jk} such that M_{jk} is the token of the structure-type M that corresponds to the token S_{jk} of the structure-type S, M_{jk} is x-morphic to S_{jk}. This will allow us to continue to talk of the relevant structure of the inclined plane model and that of the soap-box derby system and of the intended morphisms among them even if these structures are actually structure-types and not structure tokens.

III.2.7.3 No Isomorphism Holds Between these Two Relevant Structures

I will now argue that no intended isomorphism holds between the relevant structure-type of the inclined plane model and that of the soap-box derby system. Since in the standard interpretation of the model in terms of the system, the box denotes one of the racers, its position denotes the velocity of that racer, the external forces acting on it denote the "external" forces acting on the racer and so on, an intended morphism between the two structures is one that associates the box with one of the racers and the elements of T, V and R in the universe of the structure of the model with a counterpart in the universe of the structure of the system (e.g. for every $r \in R$, $f(r)=r$). The intended morphism between the two structures is an isomorphism only if, for every $t_i \in T^M$ and every $k \in I'$, $f(\mathrm{r}^M(b, t_i))=\mathrm{r}^S(f(b), f(t_i))$ and $f(\mathrm{g}^M(b, t_i, k))=\mathrm{g}^S(f(b), f(t_i), f(k))$ and $f(\mathrm{g}^M(b, t_i, k))=\mathrm{g}^S(f(b), f(t_i), f(k))$. In other words, the intended morphism is an isomorphism only if the position of the box and the forces acting on it at a certain time are identical to the position of the racer and the forces acting on it at the same time.

Now, there seems to be no need to go too much into the details of the situation to realize that this is not the case. Consider, the "external" forces acting on the racer. First of all, unlike the gravitational force on the box in the model, the gravitational force that the Earth exerts on the racer is not linear but increases as the square of the distance between the racer and the centre of mass of the earth decreases. Unlike the value of $f(\mathrm{g}^M(b, t_i, 1))$, the value of $\mathrm{g}^S(f(b), f(t_i), f(1))$, therefore, changes slightly as t_i increases and the racer goes downhill. Since, for all t, $f(\mathrm{g}^M(b, t_i, 1)) \neq \mathrm{g}^S(f(b), f(t_i), f(1))$, the intended morphism, f, cannot be an isomorphism as it does not meet condition [16.4] of definition (16).

Second, the normal force between the road and the racer is likely to be different at different times. Unlike the inclined plane, the road is likely not to be a perfectly straight slope and therefore the contact force the road exerts on the racers is likely to change as a function of time. So, unlike the value of $f(\mathrm{g}^M(b, t_i, 2))$, the value of

$g^S(f(b), f(t_i), f(2))$ is likely to differ for different values of t_i. Since, for most t_i, it is likely that $f(g^M(b, t_i, 2)) \neq g^S(f(b), f(t_i), f(2))$, the intended morphism, f, cannot be an isomorphism as it does not meet condition [16.4] of definition (16). According to definition (22), the inclined plane model, on its standard interpretation, is therefore an inexact epistemic representation of the soap-box derby system.

Third, in the inclined plane model, there are only two "external" forces acting on the racer. The value of the function $g^M(b, t_i, k)$ for every $k>2$ is the zero vector. On any of the racers in the system, on the other hand, there are a number of other "external" forces acting on the racer that have no counterpart in the model, including the air friction on the surface of the racer, the aerodynamic force on the racer, the friction between the road and the wheels of the racer, the gravitational force that any massive object in the universe exerts on the racer, from the molecules of air to distant galaxies, and so on. So, for some $k>2$ and for every $t_i \in T^M$, $f(g^M(b, t_i, k)) \neq g^S(f(b), f(t_i), f(k))$. According to definition (22), the inclined plane model, on its standard interpretation, is therefore an inexact epistemic representation of the soap-box derby system.

From the above considerations, it follows that the position of the box at each time $t_i \in T^M$ after t_0 are different from the position of the corresponding racer at that time; or, in symbols, $f(s^M(b, t_i)) \neq s^S(f(b), f(t_i))$, for all $t_i > t_0$. We can thus conclude that the inclined plane model is an inexact epistemic representation of the soap-box derby system, where an epistemic representation is *inexact* if and only if, for some $"F^V_k$, o^V_1 denotes o^T_1, …, o^V_n denotes o^T_n, o^V_i denotes o^T_i and $"F^V_k$ denotes $"F^T_k$ and $"F^V_k(o^V_1, \ldots, o^V_n) = o^V_i$ but $"F^T_k(o^T_1, \ldots, o^T_n) \neq o^T_i$.

This is far from being a peculiarity of the representation of the soap-box derby provided by the inclined plane model. Approximation and idealization characterise most if not all scientific models. For example, no model of classical mechanics instantiates a structure that is isomorphic in the intended manner to that of any real system in the sense specified above, as usually classical models do not contain a counterpart for most of the forces in the system and, even those forces for which there is a counterpart often are only an approximation of the forces in the system.

As a result of this, most models in classical mechanics are inexact epistemic representations of their target systems.

III.2.8. WEAKER MORPHISMS AND PARTIALLY FAITHFUL EPISTEMIC REPRESENTATION

In the previous section, I have argued that, whereas intended isomorphism can account for all cases of completely faithful (analytically interpreted) epistemic representation, no intended isomorphism holds between the relevant structure of those vehicles that fall short of being completely faithful epistemic representations of their targets. This is far from being a surprising result. Most sympathizers of the structuralist account realise that (intended) isomorphism is too strong a requirement in the cases of those epistemic representations that are only partially faithful and seem to think that, if a structural account of faithful epistemic representation is to deal successfully with most cases of epistemic representation and in particular with scientific models, isomorphism cannot be the morphism that holds between the relevant structure of the vehicle and that of the target. They all seem to agree however that some weaker morphism can account for partially faithful epistemic representation.

However, there is no agreement as to which of these weaker morphisms is the appropriate one. The proposed morphisms include homomorphism (see, e.g., Bartels 2006), Δ/Ψ-morphism (Swoyer 1989), partial isomorphism (see, e.g., French and Ladyman 1999 and da Costa and French 2003). In the next few sections, I will examine whether, by substituting the isomorphism in (H.2) with one of these weaker morphism that I have just mentioned, will provide us with an adequate account of partially faithful (analytically interpreted) epistemic representation.

(23) An account of partially faithful epistemic representation is *adequate* if and only if:

[23.1] it applies to all partially faithful epistemic representation (including those that are incorrect, incomplete, inexact such as the old London Underground Map and the inclined plane model and those that are completely faithful such as the new London Underground map) and

[23.2] it does not apply to any completely unfaithful epistemic representation of the target.

The intended morphism we are looking for is weak enough that it can hold between the relevant structure of the vehicle and that of the target even if the vehicle is not a completely faithful (analytically interpreted) epistemic representation of the target (even if the vehicle is an incomplete, incorrect or inexact partially faithful (analytically interpreted) epistemic representation of the target) but not so weak that it can hold between the relevant structure of the vehicle and that of the target when the vehicle is a completely unfaithful (analytically interpreted) epistemic representation of the target.

III.2.9. Homomorphism and Partially Faithful Epistemic Representation

The first morphism I will consider is homomorphism. This will give raise to what I call the *intended homomorphism account of partially faithful epistemic representation.* According to it:

(I) a vehicle is a partially faithful analytically interpreted epistemic representation of a certain target for a certain user if and only if:

(I.1) the vehicle is an analytically interpreted epistemic representation of the target for the user and

(I.2) an intended homomorphism holds between the relevant structure of the vehicle and that of the target.

I will now argue that the intended homomorphism account can account for incomplete epistemic representations that are both correct and exact, but it cannot account for partially faithful epistemic representations that are incorrect or inexact. Since the old London Underground map is both incorrect and incomplete and the inclined plane model is inexact, the intended homomorphism account cannot account for the fact that the old London Underground map and the inclined plane model are partially faithful epistemic representations of their targets. I will therefore argue first that, if an intended homomorphism holds between the relevant structure of the vehicle and that of the system, the vehicle may be an incomplete epistemic representation of the target and then that, if an intended homomorphism holds between the relevant structure of the vehicle and that of the system, the vehicle must be a correct and exact epistemic representation of the target. As usual, I will assume that the user adopts an analytic interpretation of the vehicle in terms of the target.

Incompleteness. According to (21), a vehicle is an incomplete epistemic representation of a target only if, for some $o^T_i \in A^T$, o^V_i denotes o^T_i according to the interpretation of the vehicle in terms of the target I and $o^V_i \notin A^V$ or, for some ${}^nR_k^T$, $<o^V_1, \ldots, o_n^V> \notin {}^nR_k^V$, even if $<o^T_1, \ldots, o_k^T> \in {}^nR_k^T$ and o^V_1 denotes o^T_1, …, o^V_n denotes o^T_n, ${}^nR^V_k$ denotes ${}^nR_k^T$ according to I. However, an intended homomorphism can hold between the relevant structure of the vehicle and that of the target even if , for some $o^T_i \in A^T$, o^V_i denotes o^T_i and $o^V_i \notin A^V$ or, for some ${}^nR_k^T$, $<o^V_1, \ldots, o^V_n> \notin {}^nR_k^V$, even if $<o^T_1, \ldots, o_k^T> \in {}^nR_k^T$ and o^V_1 denotes o^T_1, …, o^V_n denotes o^T_n, ${}^nR^V_k$ denotes ${}^nR_k^T$. For an intended homomorphism to hold between the relevant structure of the vehicle, it is necessary that, for every object o^V_i in the universe of the relevant structure of the vehicle, the object denoted by o^V_i, o^T_i, is in the universe of the relevant structure of the target. However, an intended homomorphism can hold even if, for some object o^T_i in the universe of the relevant structure of the target, there is no object in the universe of the relevant structure the vehicle that denotes o^T_i. Analogously, for an intended homomorphism to hold between the relevant structure of the vehicle, it is necessary that if ${}^nR^V_k$ denotes ${}^nR_k^T$ and $<o^V_1, \ldots, o_n^V> \in {}^nR_k^V$, $<o^T_1, \ldots, o_k^T> \in {}^nR_k^T$, however it is not necessary that, if $<o^T_1, \ldots, o_n^T> \in {}^nR_k^T$, $<o^V_1, \ldots, o^V_k> \in {}^nR^V_k$. So, if an intended homomorphism holds between the relevant structure of the vehicle and that of the target, the vehicle may well be an incomplete epistemic representation of the target.

Correctness. If an intended homomorphism holds between the relevant structure of the vehicle and the relevant structure of the target, the vehicle cannot be an incorrect epistemic representation of the target. According to definition (20), if the vehicle was an incorrect epistemic representation of a target, it would be the case that, for some $o^{V}_{i} \in A^{V}$, o^{V}_{i} denotes o^{T}_{i} and $o^{T}_{i} \notin A^{T}$ or, for some ${}^{n}R^{V}_{k}$, $<o^{T}_{1}, \ldots, o_{n}{}^{T}> \notin {}^{n}R_{k}{}^{T}$, even if $<o^{V}_{1}, \ldots, o^{V}_{k}> \in {}^{n}R^{V}_{k}$ and o^{V}_{1} denotes $o^{T}_{1}, \ldots, o^{V}_{n}$ denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R_{k}{}^{T}$. If this was the case, however no intended homomorphism could hold between the relevant structure of the vehicle and that of the target because, according to definitions (15) and (19), if an intended homomorphism held between the relevant structure of the vehicle and the target, then every object $o^{V}_{i} \in A^{V}$ would be associated with the object $o^{T}_{i} \in A^{T}$ denoted by it and, for every ${}^{n}R_{k}{}^{T}$, if $<o^{V}_{1}, \ldots, o^{V}_{k}> \in {}^{n}R^{V}_{k}$, then $<o^{T}_{1}, \ldots, o_{k}{}^{T}> \in {}^{n}R_{k}{}^{T}$. So, if an intended homomorphism holds between the relevant structure of the vehicle and that of the target, the vehicle must be a correct epistemic representation of the target.

Exactness. If an intended homomorphism holds between the relevant structure of the vehicle and the relevant structure of the target, the vehicle cannot be an inexact epistemic representation of the target. If the vehicle was an inexact representation of the target, it would be the case that, for some ${}^{n}F^{V}_{k}$, o^{V}_{1} denotes o^{T}_{1}, $\ldots$, o^{V}_{n} denotes o^{T}_{n}, and ${}^{n}F^{V}_{k}$ denotes ${}^{n}F^{T}_{k}$ and $f({}^{n}F^{V}_{k}(o^{V}_{1}, \ldots, o^{V}_{n}) \neq {}^{n}F^{T}_{k}(f(o^{V}_{1}), \ldots, f(o^{V}_{n}))$. If this was the case, however no intended homomorphism could hold between the relevant structure of the vehicle and that of the target because, according to definitions (15) and (19), if an intended homomorphism held between the relevant structure of the vehicle and that of the target, then for every ${}^{n}F^{V}_{k}$, if o^{V}_{1} denotes $o^{T}_{1}, \ldots, o^{V}_{n}$ denotes o^{T}_{n}, and ${}^{n}F^{V}_{k}$ denotes ${}^{n}F^{T}_{k}$, $f({}^{n}F^{V}_{k}(o^{V}_{1}, \ldots, o^{V}_{n})$ would have to be equal to ${}^{n}F^{T}_{k}(f(o^{V}_{1}), \ldots, f(o^{V}_{n}))$. So, if an intended homomorphism holds between the relevant structure of the vehicle and that of the target, the vehicle must be an exact epistemic representation of the target.

The situation is analogous in the case of what we could call the *intended inverse homomorphism account of partially faithful epistemic representation.* Structure A is inversely homomorphic to structure **B** if and only if structure **B** is homomorphic to structure A. According to the intended inverse homomorphism account:

(J) a vehicle is a partially faithful, analytically interpreted epistemic representation of a certain target for a certain user if and only if:

(J.1) the vehicle is an analytically interpreted epistemic representation of the target for the user and

(J.2) an intended homomorphism holds between the relevant structure of the vehicle and that of the target.

The difference here is that the intended homomorphism is between the relevant structure of the target and that of the vehicle not the reverse and, since homomorphism is not a symmetric relation, the intended inverse homomorphism account of partially faithful epistemic representation is different from the intended homomorphism account.

I will now argue that the inverse homomorphism account of partially faithful epistemic representation can account for incorrect epistemic representations that are incorrect but are both complete and exact.

Incorrectness. According to definition (20), a vehicle is an incorrect epistemic representation of a target only if, for some $o^{V}_{i} \in A^{V}$, o^{V}_{i} denotes o^{T}_{i} and $o^{T}_{i} \notin A^{T}$ or, for some ${}^{n}R^{V}_{k}$, o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$ and $<o^{V}_{1}, \ldots, o^{V}_{n}> \in {}^{n}R^{V}_{k}$ but $<o^{T}_{1}, \ldots, o_{n}^{T}> \notin {}^{n}R^{T}_{k}$. However, an intended homomorphism can hold between the relevant structure of the target and that of the vehicle even if, for some $o^{V}_{i} \in A^{V}$, o^{V}_{i} denotes o^{T}_{i} and $o^{T}_{i} \notin A^{T}$. or, for some ${}^{n}R^{V}_{k}$, o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$ and $<o^{V}_{1}, \ldots, o^{V}_{k}> \in {}^{n}R^{V}_{k}$, but $<o_{1}^{T}, \ldots, o^{T}_{n}> \notin {}^{n}R^{T}_{k}$. For an intended homomorphism to hold between the relevant structure of the target and that of the vehicle, it is necessary that, for every object, o^{T}_{i}, in the universe of the relevant structure of the target, the object that denotes o^{T}_{i}, o^{V}_{i}, is in the universe relevant structure of the vehicle. However, there can be objects in the universe of the relevant structure of the vehicle even if the objects that denote them are not in the relevant structure of the target. Analogously, for an intended homomorphism to hold between the relevant structure of the target and that of the vehicle, it is necessary that if ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$ and $<o^{T}_{1}, \ldots, o^{T}_{n}> \in {}^{n}R^{T}_{k}$, $<o^{V}_{1}, \ldots, o^{V}_{n}> \in {}^{n}R^{V}_{k}$, however it is not necessary that if $<o^{V}_{1}, \ldots, o^{V}_{n}> \in {}^{n}R^{V}_{k}$, $<o^{T}_{1}, \ldots, o^{T}_{n}> \in {}^{n}R^{T}_{k}$. So, if an intended homomorphism holds between the relevant structure of the target and that of the vehicle, the vehicle can be an incorrect representation of the target.

Completeness. If an intended homomorphism holds between the relevant structure of the target and the relevant structure of the vehicle, the vehicle cannot be an incomplete epistemic representation of the target. If the vehicle was an incomplete epistemic representation of a target, it would be the case that, for some $o^{T}_{i} \in A^{T}$, o^{V}_{i} denotes o^{T}_{i} and $o^{V}_{i} \notin A^{V}$ or, for some ${}^{n}R^{T}_{k}$, o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$, and $<o^{T}_{1}, \ldots, o_{k}^{T}> \in {}^{n}R^{T}_{k}$, but $<o^{V}_{1}, \ldots, o_{n}^{V}> \notin {}^{n}R^{V}_{k}$. If this were the case, however no intended homomorphism could hold between the relevant structure of the target and that of the vehicle because if an intended homomorphism held between the relevant structure of the target and that of the vehicle, then every object $o^{T}_{i} \in A^{T}$ would be associated with the object $o^{V}_{i} \in A^{V}$ that denotes it and, for every ${}^{n}R^{T}_{k}$, if ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$ and $<o^{T}_{1}, \ldots, o^{T}_{n}> \in {}^{n}R^{T}_{k}$, then it would have to be the case that $<o^{V}_{1}, \ldots, o^{V}_{n}> \in {}^{n}R^{V}_{k}$. So, if an intended homomorphism holds between the relevant structure of the vehicle and that of the target, the vehicle must be a complete epistemic representation of the target.

Exactness. If an intended homomorphism holds between the relevant structure of the target and that of the vehicle, the vehicle cannot be an inexact epistemic representation of the target. If the vehicle was an incorrect representation of the target, it would be the case that, for some ${}^{n}F^{V}_{k}$, o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, and ${}^{n}F^{V}_{k}$ denotes ${}^{n}F^{T}_{k}$ and $f({}^{n}F^{V}_{k}(o^{V}_{1}, \ldots, o^{V}_{n}) \neq {}^{n}F^{T}_{k}(f(o^{V}_{1}), \ldots, f(o^{V}_{n}))$. If this were the case, however no intended homomorphism could hold between the relevant structure of the target and that of the vehicle because, if an intended homomorphism held between the relevant structure of the target and that of the vehicle, then for every ${}^{n}F^{T}_{k}$, if o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, and ${}^{n}F^{V}_{k}$ denotes ${}^{n}F^{T}_{k}$, it would have to be the case that $f({}^{n}F^{T}_{k}(o^{T}_{1}, \ldots, o^{T}_{n}) = {}^{n}F^{V}_{k}(f(o^{T}_{1}), \ldots, f(o^{T}_{n}))$. So, if an intended homomorphism holds between the relevant structure of the target and that of the vehicle, the vehicle must be an exact epistemic representation of the target.

Each of the two accounts of partially faithful epistemic representation that I have examined in this subsection can only account for very specific kinds of partially faithful epistemic representation. However, neither can account for partially faithful epistemic representations that are both incorrect and incomplete like the old London Underground map or for partially faithful epistemic representations that are inexact like the inclined plane model. Both accounts are therefore inadequate.

III.2.10. Δ/Ψ-MORPHISM AND PARTIALLY FAITHFUL EPISTEMIC REPRESENTATION

The notion of Δ/Ψ-morphism was introduced by Chris Swoyer (1991). Unlike the other morphisms that we have considered so far, a Δ/Ψ-morphism is not a relation between set-theoretic structures but a relation between what Swoyer calls *intensional relational system* (*IRS*). An *IRS* is an ordered quadruple $S = \langle A^S, {}^{I}R^S, {}^{II}R^S, ext\rangle$, where A^S is a set of individuals, ${}^{I}R^S$ is a set of first-order relations, ${}^{II}R^S$ is a set of second-order relations (where, as usual, properties are construed as unary relations), *ext* is a function that assigns to a relation its extension and at most one of A^S and ${}^{II}R^S$ is empty. Unlike the relations in a set-theoretic structure, the relations in an intensional relational system are not construed extensionally. This means that two relations may have the same extension without being the same relation ($ext(R_i^T)=ext(R_j^T)$ does not imply that $R_i^T=R_j^T$). For example, if all and only the London Underground stations on the Victoria line have three escalators, the properties of *being on the Victoria Line* and that of *having three escalators* have the same extension but they are not the same property.

(24) A function, *f*, from $A^T \cup {}^{I}R^T \cup {}^{II}R^T$ to $A^V \cup {}^{I}R^V \cup {}^{II}R^V$ is a *Δ/Ψ-morphism* if and only if:

[24.1] Δ and Ψ are subsets of ${}^{I}R^T \cup {}^{II}R^T$ and at least Δ is not empty.

[24.2] For all ${}^{I}R_i^T$ and ${}^{II}R_i^T$ in Δ, if $<a_1^T, \ldots, a_k^T> \in ext(R_i^T)$, then $<f(a_1^T), \ldots, f(a_k^T)> \in ext(f(R_i^T))$ (*f preserves* all relations in Δ).

[24.3] For all ${}^{I}R^T$ and ${}^{II}R^T$ in Ψ, if $<f(a_1^T), \ldots, f(a_k^T)> \in ext(f(R_i^T))$, then $<a_1^T, \ldots, a_k^T> \in ext(R_i^T)$ (*f counter-preserves* all relations in Ψ).

The *intended Δ/Ψ-morphism account of partially faithful epistemic representation* maintains that:

(K) a vehicle is a partially faithful (analytically interpreted) epistemic representation of a certain target for a user if and only if:

(K.1) the vehicle is an analytically interpreted epistemic representation of the target for that user and

(K.2) an intended Δ/Ψ-morphism holds between the relevant IRS of the target and that of the vehicle.

In this subsection, I will argue that the intended Δ/Ψ-morphism is not a successful accounting of partially faithful epistemic representation. A first, minor problem is that Δ/Ψ-morphism does not allow a certain kind of incomplete representation to be partially faithful epistemic representation. A Δ/Ψ-morphism is a function from $A^T \cup {}^{I}R^T \cup {}^{II}R^T$ to $A^V \cup {}^{I}R^V \cup {}^{II}R^V$. This means that the intended Δ/Ψ-morphism associates every object in A^T with the object that denotes it in A^V. However, an epistemic representation can still be partially faithful even if for some objects in A^T, the objects that denote them are not in A^V. For example, some stations on the London Underground network have no counterpart on the old London Underground map, but nevertheless we consider the latter a partially faithful epistemic representation of the former. This problem can easily be avoided by modifying the definition of Δ/Ψ-morphism so that a Δ/Ψ-morphism is a function, *f*, from *a non-empty subset of* the universe of the relevant IRS of the target.

A second, much more serious problem arises from the very notions of preservation and counter-preservation of a relation (from conditions [24.2] and

[24.3] of the definition of a Δ/Ψ-morphism). A relation among objects in the universe of the relevant IRS of the target (e.g. the relation *being connected by Metropolitan Line trains* in the London Underground network) is preserved if and only if, for any two stations, if those stations are connected by Metropolitan line trains, then the circles or tabs that denote those stations are connected by a maroon line, it is counter-preserved if and only if, for any two circles or tabs, if those circles or tabs are connected by a maroon line, then the stations they denote are connected by Metropolitan Line trains. So, if the relation *being connected by Metropolitan Line trains* is preserved, it is sound to infer that two stations are not connected by Metropolitan Line trains if the circles or tabs that denote them are not connected by a maroon line; if the relation is counter-preserved, it is sound to infer that two stations are connected by Metropolitan Line trains if the circles or tabs that denote them are connected by a maroon line.

According to the intended Δ/Ψ-morphism account, a vehicle is a partially faithful epistemic representation of a target only if some properties of or relations among objects in the target are counter-preserved. The rationale behind this requirement is that, if the Δ/Ψ-morphism holds, then it is possible to explain why some of the inferences performed by the user are sound—they are sound because if the relation holds among the objects in the vehicle, the relation denoted by it also holds among the objects in the target that are denoted by those objects in the vehicle—and therefore why the epistemic representation in question is a partially faithful one.

However, the requirement is too strong—an epistemic representation may be partially faithful even if no property or relation in the target is counter-preserved. On the old London Underground map, for example, the circles labelled 'Aldgate' and 'Hammersmith' are connected by a maroon line but the corresponding stations are not connected to by Metropolitan Line trains. Therefore the relation *being connected by Metropolitan Line trains* is not counter-preserved by any intended Δ/Ψ-morphism between the relevant IRS of the map and that of the network. Nevertheless, one can perform many sound inferences from the fact that circles or tabs are connected by a maroon line to the fact that the stations denoted by them are connected by Metropolitan line trains. So, one can perform sound inferences from the fact that a certain relation holds among certain objects in the vehicle to the fact that the relation denoted by it holds among the objects in the target denoted by them even if the relation is not counter-preserved. Therefore, even if no relation among objects in the target is counter-preserved, the vehicle may still be a partially faithful epistemic representation of the target if, for some objects in the vehicle, from the fact that a certain relation, *R*, holds among them, it is sound to infer that the relation denoted by *R* holds between the objects in the target that are denoted by those objects.

Here it is important to note that an account of partially faithful epistemic representation is meant to identify by virtue of what an epistemic representation is a faithful one not by virtue of what the user knows or believes it is a faithful one. A user may not be able to determine which of the inferences from the fact that circles or tabs are connected by a maroon line to the fact that the stations denoted by them are connected by Metropolitan line trains are sound and which are not, but the representation is still faithful to some degree if some of the inferences from it to the target are sound.

This intended Δ/Ψ-morphism conception is not even successful in accounting for one of the cases that Swoyer seems to have in mind when introducing the notions of preservation and counter-preservation—the case of topographical maps. Swoyer correctly points out:

> [...] it is a basic geometrical fact that a two-dimensional projection of a sphere cannot depict all its features without distortion, so when we use flat maps to represent the Earth, something has to give. For sixteenth-century mariners, concerned to convert lines of constant compass bearing (rhumb lines) into straight lines on their maps, Mercator's projection, which misrepresents scale, offered the best compromise; for other purposes equal area maps, which

accurately represent scale but not shape, are preferable (Swoyer 1989, 470).

Swoyer is right in claiming that some maps use projections that counter-preserve some properties or relations of the geographical area represented. For example, the Polar Azimuthal projection counter-preserves the distance of every point from the North Pole, but it does so at the cost of distorting shapes and areas.[23] And this is a far from being a unique case. Projections that counter-preserve one property or relation do so at the cost of extreme distortion of other properties or relations. However, it is exactly for this reason that many of the projections that are most commonly used, such as Lambert's Conformal Conic projection, do not counter-preserve any property but rather attempt to minimize the distortion of as many properties or relations as possible (see, for example, (Fisher and Miller 1944)). Since maps based on these projections do not counter-preserve any property or relation, the intended Δ/Ψ-morphism conception cannot account for the fact that they are partially faithful epistemic representation of their targets.

As the case of the maps based on the Lambert's Conformal Conic projection shows, a representation may be partially faithful even if none of the relations among objects in the universe of the target is counter-preserved by any intended Δ/Ψ-morphism. The intended Δ/Ψ-morphism account of partially faithful epistemic representation is therefore inadequate as an account of partially faithful epistemic representation.

[23] See (Fisher and Miller 1994), which is an excellent introduction to topographical map projections, or the more technical (Monmonier 1977) and (Maling 1992).

III.2.11. PARTIAL ISOMORPHISM AND PARTIALLY FAITHFUL EPISTEMIC REPRESENTATION

III.2.11.1 The Partial Isomorphism Conception of Partially Faithful Epistemic Representation

According to the intended partial isomorphism conception of partially faithful epistemic representation,

(L) a vehicle is a partially faithful, analytically interpreted epistemic representation of a certain target for a certain user if and only if:

(L.1) the vehicle is an analytically interpreted epistemic representation of the target,

(L.2) the user adopts an analytic interpretation of the vehicle in terms of the target, and

(L.3) a (non-vacuous) intended partial isomorphism holds between the relevant structure of the vehicle and that of the target.

Two structures, **A** and **B**, are *partially isomorphic* if and only if there are two partial substructures of **A** and **B** that are isomorphic. The best way to introduce the notion of partial substructure is to introduce the notion of a *substructure* and the notion of a *partial* structure first and only then introduce the notion of a *partial substructure.*

(25) A (total) structure **B** is *a substructure of* a (total) structure **A** if and only if:

[25.1] The universe of **B**, A^{B}, is a subset of the universe of **A**, A^{A},

[25.2] For every n-ary relation, R_i^{nA}, in **A**, there is an n-ary relation in **B** such that $R_i^{nB} = R_i^{nA} \cap (A^{B})^n$ (more informally, for every relation in **A**, there is a relation in **B** whose extension is the subset of the relation in **A** that only contains the n-tuples of elements that are in the universe of **B**).

[25.3] For every n-ary function, F_j^{nA}, in A, there is an n-ary function in B such, that, if $dom(F_j^{nA})$ is the domain of F_j^{nA} and $codom(F_j^{nA})$ is the codomain of F_j^{nA}, $dom(F_j^{nB}) = dom(F_j^{nA}) \cap (A^B)^n$ and $codom(F_j^{nB}) = codom(F_j^{nA}) \cap (A^B)^n$ (more informally, for every function in A, there is a function in B whose domain is the subset of the universe of A that only contains the n-tuples of elements that are in the universe of B) (cf. van Dalen 1991).

A *partial structure* is an n-tuple $P = \langle A^P, R_1^{mP}, \ldots, R_j^{oP}, F_1^{pP}, \ldots, F_k^{rP} \rangle$, which is defined analogously to a total structure except for the fact that $R_1^{mP}, \ldots, R_j^{oP}$ are *partial relations* on A^P and $F_1^{pP}, \ldots, F_k^{rP}$ are *partial functions.* A partial m-ary relation, R_i^{mP}, is a triple whose first element is the set of m-tuples of elements of A^P, $sat(R_i^{mP})$, that satisfy the relation R_i^{mP}, the second is the set of m-tuples of elements of A^P, $dissat(R_i^{mP})$, that do not satisfy the relation R_i^{mP}, the third element is the set of elements for which it is not specified whether they satisfy nor do not satisfy the relation R_i^{mP} $indet(R_i^{mP})$. A total relation can be thus seen as a limit case of a partial relation whose third component is the empty set. A *partial function* is a function that may not assign any value to some arguments within its domain.[24]

We can now introduce the notion of partial substructure.

(26) A partial structure **B** is *a partial substructure of* the total structure **A** if and only if:

[26.1] The universe of B, A^B, is a subset of the universe of A, A^A (so, for every $o^B_i \in A^B$, $o^B_i = o^A_i$)

[24] It is worth noting that my definition of a partial structure differs from the notion employed by da Costa, French and their collaborators in one aspect, which will turn out to be crucial to my conception. Da Costa and French restrict their analysis to structures that contain relations but not functions. Obviously, this is not a problem in itself. Any n-ary function can be recasted as a $(n+1)$-ary relation whose *relata* are the arguments of the function and its value. In fact, a partial unary function, $f(x)$, that does not assign a value to a certain argument, a, in its domain can be seen as a relation that is undefined for any ordered couple whose first component is a and whose second component is an element of the co-domain of the function.

[26.2] For every *n*-ary relation, ${}^{n}R^{A}_{i}$, in **A**, there is an *n*-ary relation in **B** such that $sat({}^{n}R^{B}_{i}) \subseteq ({}^{n}R^{A}_{i} \cap (A^{B})^{n})$ and $dissat({}^{n}R^{B}_{i}) \cap {}^{n}R^{A}_{i} = \varnothing$ (informally, a certain *n*-tuple of elements of the universe of **B** belongs to the set of *n*-tuples that satisfy a certain relation in **B** *only if* it is in the extension of the corresponding relation in **A** and it belongs to the set of *n*-tuples that do not satisfy that *only if* it is not in the extension of the corresponding relation in **A**.).[25]

[26.3] For every *n*-ary function, ${}^{n}F^{A}_{i}$, in A, there is an *n*-ary function in B such, that, if $dom({}^{n}F^{A}_{i})$ is the domain of ${}^{n}F^{A}_{i}$ and $codom({}^{n}F^{A}_{i})$ is the codomain of ${}^{n}F^{A}_{i}$, $dom({}^{n}F^{B}_{i}) = (dom({}^{n}F^{A}_{i}) \cap (A^{B})^{n})$ and $codom({}^{n}F^{B}_{i}) = codom({}^{n}F^{A}_{i}) \cap (A^{B})^{n}$ (informally, for every function in A, there is a function in **B** whose domain is the subset of the domain of the function in A that only contains the *n*-tuples of elements that are in the universe of **B** and whose codomain is the subset of the domain of the function in A that only contains the *n*-tuples of elements that are in the universe of **B**).

Since the isomorphism between the partial substructures of the relevant structure of the vehicle, **V**, and that of the target, **T**, is an *intended* isomorphism, all of the intendedly isomorphic partial substructures—i.e. all of the partial

[25] Let me note that this is markedly different from how French and his collaborators define these concepts. For example, French and Ladyman (1999) claim that a certain *n*-tuple of elements of the universe of B belongs to the set of *n*-tuples that satisfy a certain relation in B *if* and only if it is in the extension of the corresponding relation in A; it belongs to the set of *n*-tuples that do not satisfy that relation *if* and only if it is not in the extension of the corresponding relation in A. However, this cannot possibly be what French and Ladyman actually have in mind because, according to this definition, the set $indet({}^{n}R_{i}^{mP})$ would always be empty and therefore the set of partial substructures of A would be identical to that of the substructures of A. In order for B to be a (genuine) *partial* substructure of A it must be the case that $indet({}^{n}R_{i}^{P})$ is not empty. This is accomplished by the definition I have put forward here but not by the one that French and his collaborators consider.

substructures of V and T such that an intended isomorphism holds between them—must meet the following conditions:

(i) If o_i^{V} denotes o_i^{T}, $o^{V^*}_i \in A^{V^*}$ and $o_i^{T^*} \in A^{T^*}$ only if $o_i^{V} \in A^{V}$ and $o_i^{T} \in A^{T}$.

(ii) If o_1^{V} denotes o_1^{T}, ..., o_n^{V} denotes o_n^{T}, ${}^{n}R_k^{V}$ denotes ${}^{n}R_k^{T}$, $o_1^{V^*} \in A^{V^*}$, ..., $o_n^{V^*} \in A^{V^*}$, then:

[26.4] $<o_1^{V^*}, \ldots, o_n^{V^*}> \in sat({}^{n}R_k^{V^*})$ and $<o_1^{T^*}, \ldots, o_n^{T^*}> \in sat({}^{n}R_k^{T^*})$ only if $<o_1^{V}, \ldots, o_n^{V}> \in {}^{n}R_k^{V}$ and $<o_1^{T}, \ldots, o_n^{T}> \in {}^{n}R_k^{T}$,

[26.5] $<o_1^{V^*}, \ldots, o^{V^*}_n> \in dissat({}^{n}R^{V^*}_k)$ and $<o_1^{T^*}, \ldots, o_n^{T^*}> \in sat({}^{n}R_k^{T^*})$ only if $<o_1^{V}, \ldots, o_n^{V}> \notin {}^{n}R_k^{V}$ and $<o^{T}_1, \ldots, o^{T}_n> \notin {}^{n}R_k^{T}$, and

[26.6] $<o_1^{V^*}, \ldots, o_n^{V^*}> \in indet({}^{n}R_k^{V^*})$ and $<o_1^{V^*}, \ldots, o_n^{V^*}> \in dissat({}^{n}R_k^{V^*})$ only if either $<o_1^{V}, \ldots, o_n^{V}> \in {}^{n}R_k^{V}$ and $<o_1^{T}, \ldots, o_n^{T}> \notin {}^{n}R_k^{T}$ or $<o_1^{V}, \ldots, o_n^{V}> \notin {}^{n}R_k^{V}$ and $<o_1^{T}, \ldots, o_n^{T}> \in {}^{n}R_k^{T}$.

(iii) If o_1^{V} denotes o_1^{T}, ..., o_n^{V} denotes o_n^{T}, o_i^{V} denotes o_i^{T}, ${}^{n}F_k^{V}$ denotes ${}^{n}F_k^{T}$, $o_1^{V^*} \in A^{V^*}$, ..., $o_n^{V^*} \in A^{V^*}$, $o_1^{T^*} \in A^{T^*}$, ..., and $o_n^{T^*} \in A^{T^*}$, then

[26.7] ${}^{n}F_k^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*}) = o_i^{V^*}$ and ${}^{n}F_k^{T^*}(o_1^{T^*}, \ldots, o_n^{T^*}) = o_i^{T^*}$ only if ${}^{n}F_k^{V}(o_1^{V}, \ldots, o_n^{V}) = o_i^{V}$ and ${}^{n}F_k^{T}(o_1^{T}, \ldots, o^{T}_n) = o_i^{T}$,

[26.8] ${}^{n}F^{V^*}_k(o^{V^*}_1, \ldots, o_n^{V^*})$ and ${}^{n}F^{T^*}_k(o_1^{T^*}, \ldots, o_n^{T^*})$ are indeterminate only if ${}^{n}F_k^{V}(o_1^{V}, \ldots, o_n^{V}) \neq o_i^{V}$ or ${}^{n}F_k^{T}(o_1^{T}, \ldots, o_n^{T}) \neq o_i^{T}$.

Two particularly important intendedly isomorphic partial substructures of the relevant structures V and T are, V* and T*, which are what I will call the *maximal intendedly isomorphic partial substructures.*

(27) If V and T are the relevant structure of, respectively, a vehicle, *V*, and a target, *T*, relative to a certain interpretation of the vehicle in terms of the target *I*, V* and T*, are the maximal intendedly isomorphic partial substructures of respectively V and T if and only if:

[27.1] V* is a partial substructure of V;

[27.2] T* is a partial substructure of T,

[27.3] If, according to I, o_i^{V} denotes o_i^{T}, $o^{V^*}_i \in A^{V^*}$ and $o_i^{T^*} \in A^{T^*}$ if and only if $o_i^{V} \in A^{V}$ and $o_i^{T} \in A^{T}$.

[27.4] If, according to I, o_1^{V} denotes o_1^{T}, ..., o_n^{V} denotes o_n^{T}, $^{n}R_k^{V}$ denotes $^{n}R_k^{T}$, $o_1^{V^*} \in A^{V^*}$, ..., $o_n^{V^*} \in A^{V^*}$, then:

[27.4.1] $<o_1^{V^*}, \ldots, o_n^{V^*}> \in sat(^{n}R_k^{V^*})$ and $<o_1^{T^*}, \ldots, o_n^{T^*}> \in sat(^{n}R_k^{T^*})$ if and only if $<o_1^{V}, \ldots, o_n^{V}> \in {}^{n}R_k^{V}$ and $<o_1^{T}, \ldots, o_n^{T}> \in {}^{n}R_k^{T}$,

[27.4.2] $<o_1^{V^*}, \ldots, o_n^{V^*}> \in dissat(^{n}R_k^{V^*})$ and $<o_1^{T^*}, \ldots, o_n^{T^*}> \in sat(^{n}R_k^{T^*})$ if and only if $<o_1^{V}, \ldots, o_n^{V}> \notin {}^{n}R_k^{V}$ and $<o_1^{T}, \ldots, o_n^{T}> \notin {}^{n}R_k^{T}$, and

[27.4.3] $<o_1^{V^*}, \ldots, o_n^{V^*}> \in indet(^{n}R_k^{V^*})$ and $<o_1^{V^*}, \ldots, o_n^{V^*}> \in dissat(^{n}R_k^{V^*})$ if and only if either $<o_1^{V}, \ldots, o_n^{V}> \in {}^{n}R_k^{V}$ and $<o_1^{T}, \ldots, o_n^{T}> \notin {}^{n}R_k^{T}$ or $<o_1^{V}, \ldots, o_n^{V}> \notin {}^{n}R_k^{V}$ and $<o_1^{T}, \ldots, o_n^{T}> \in {}^{n}R_k^{T}$.

[27.5] If, according to I, o_1^{V} denotes o_1^{T}, ..., o_n^{V} denotes o_n^{T}, o_i^{V} denotes o_i^{T}, $^{n}F_k^{V}$ denotes $^{n}F_k^{T}$, $o_1^{V^*} \in A^{V^*}$, ..., $o_n^{V^*} \in A^{V^*}$, $o_1^{T^*} \in A^{T^*}$, ..., and $o_n^{T^*} \in A^{T^*}$, then

[27.5.1] $^{n}F_k^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*}) = o_j^{V^*}$ and $^{n}F_k^{T^*}(o_1^{T^*}, \ldots, o_n^{T^*}) = o_i^{T^*}$ if and only if $^{n}F_k^{V}(o_1^{V}, \ldots, o_n^{V}) = o_i^{V}$ and $^{n}F_k^{T}(o_1^{T}, \ldots, o_n^{T}) = o_i^{T}$,

[27.5.2] $^{n}F_k^{V^*}(o_1^{V^*}, \ldots, o^{V^*}{}_n)$ and $^{n}F_k^{T^*}(o^{T^*}{}_1, \ldots, o^{T^*}{}_n)$ are indeterminate if and only if $^{n}F_k^{V}(o_1^{V}, \ldots, o_n^{V}) \neq o_i^{V}$ or $^{n}F_k^{T}(o^{T}{}_1, \ldots, o^{T}{}_n) \neq o^{T}{}_i$.

Clearly V* and T* are intendedly isomorphic. Moreover, V* and T* are maximal in the sense that any intendedly partial substructures of V and T are either identical with V* and T* or are proper partial substructures of V* and T*. Let me illustrate this with a concrete example. Consider the case of the old London

Underground map, which is both incomplete and incorrect. First of all, the map contains circles and tabs that do not denote any station on the network (e.g. the tab labelled 'Dover Street') and the network contains stations that are not denoted by any station in the map (e.g. Bethnal Green station). The maximally intendedly isomorphic partial substructures of the relevant structures of the map and the network will thus be those substructures whose universes contain respectively *all and only* those circles and tabs that denote stations and *all and only* those stations that are denoted by circles or tabs on the map. So, whereas the universe of the relevant structure of the vehicle contains the tab labelled 'Dover Street' and that of the relevant structure of the target contains Bethnal Green station, the universe of the maximally intendedly isomorphic partial substructure of the vehicle does not contain the tab labelled 'Dover Street' and that of the maximally intendedly isomorphic partial substructure of the target does not contain Bethnal Green station.

Second, some circles and tabs have certain properties and are in certain relations, even if the stations denoted by those circles and tabs do not have the corresponding properties or do not stand in the corresponding relation. For example, the circles labelled 'Aldgate' and 'Hammersmith' are connected by a maroon line even if the stations denoted by those circles, Aldgate station and Hammersmith station, are not connected by Metropolitan Line trains. Whereas, in the relevant structure of the target, the couple <tab labelled 'Aldgate', circle labelled 'Hammersmith'> belongs to the set of all couples of circles or tabs that are connected by a maroon line and the couple <Aldgate, Hammersmith> does *not* belong to the set of couples that are connected by Metropolitan Line trains, in the maximally intendedly isomorphic partial substructures, the couple <tab labelled 'Aldgate', circle labelled 'Hammersmith'> belongs to the set of couples for which it is undetermined whether they satisfy nor do not satisfy the relation *being connected by a maroon line* and the couple <Aldgate, Hammersmith> belongs to the set of couples of stations for which it is undetermined whether they satisfy nor do not satisfy *being connected by Metropolitan Line trains.* Analogously, some stations have certain properties and are in certain relations, despite the fact that the circles and tabs on the old map that denote those stations do not have the corresponding properties or do not stand in the corresponding relation.

To put it figuratively, the maximally intendedly isomorphic partial substructures of the relevant structures contain all and only those bits of each relevant structure that have an intended counterpart in the other structure. That is, an object in the universe of the relevant structure of the vehicle is contained in the universe of the maximally intendedly isomorphic partial substructure of the relevant structure of the vehicle only if the object denoted by it according to the interpretation adopted by the user is contained in the universe of the relevant structure of the target and vice versa—an object in the universe of the relevant structure of the target is contained in the universe of the maximally intendedly isomorphic partial substructure of the relevant structure of the target only if the object that denotes it according to the interpretation adopted by the user is contained in the universe of the relevant structure of the vehicle. A certain n-tuple that satisfies a certain relation in the relevant structure of the vehicle is in the set of n-tuples that satisfy the corresponding relation in the maximally intendedly isomorphic partial substructure of the relevant structure of the vehicle only if the n-tuple of objects in the relevant structure of the target denoted by them are in the relation denoted by that relation and so on.

III.2.11.2 The Partial Success of the Intended Partial Isomorphism Account

Leaving aside the crucial issue of inexactness and the proviso that the partial isomorphism be non-vacuous until the next section, I will now argue that the intended partial isomorphism account of partially faithful epistemic representation can account successfully for analytically interpreted epistemic representations that are incomplete, incorrect or both.

Incorrectness. A vehicle is an incorrect epistemic representation of a target if and only if, for some $o^V_i \in A^V$, o^V_i denotes o^T_i and $o^T_i \notin A^T$ or, for some ${}^nR^V_k$, o^V_1 denotes o^T_1, …, o^V_n denotes o^T_n, ${}^nR^V_k$ denotes ${}^nR^T_k$ and $<o^V_1, \ldots, o^V_n> \in {}^nR^V_k$. but $<o^T_1, \ldots, o_n^T> \notin {}^nR^T_k$. However, an intended partial isomorphism can hold between the relevant structure of the vehicle and that of the target even if, for some $o^{V}_{i} \in A^{V}$, o^{V}_{i} denotes o^{T}_{i} and $o^{T}_{i} \notin A^{T}$ or, for some ${}^{n}R^{V}_{k}$, o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$ and $<o^{V}_{1}, \ldots, o^{V}_{k}> \in {}^{n}R^{V}_{k}$, but $<o^{T}_{1}, \ldots, o^{T}_{n}> \notin {}^{n}R^{T}_{k}$. For an intended partial isomorphism to hold between the relevant structure of the vehicle and that of the target, an object can be an element of the universe of the relevant structure of the vehicle even if the objects that denotes it is not in the relevant structure of the target—the first object will simply not be in the universe of the maximal intendedly isomorphic *partial sub*structure of the vehicle. Analogously, for an intended partial isomorphism to hold between the relevant structure of the vehicle and that of the target, it is not necessary that, if o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$, and $<o^{V}_{1}, \ldots, o^{V}_{n}> \in {}^{n}R^{V}_{k}$, then $<o^{T}_{1}, \ldots, o^{T}_{n}> \in {}^{n}R^{T}_{k}$. If $<o^{V}_{1}, \ldots, o^{V}_{n}> \in {}^{n}R^{V}_{k}$ and $<o^{T}_{1}, \ldots, o^{T}_{n}> \notin {}^{n}R^{T}_{k}$, it will simply be the case that $<o^{V*}_{1}, \ldots, o^{V*}_{n}> \in indet({}^{n}R^{V*}_{k})$ and $<o^{T*}_{1}, \ldots, o^{T*}_{n}> \in indet({}^{n}R^{T*}_{k})$. So, if an intended partial isomorphism holds between the relevant structure of the target and that of the vehicle, the vehicle can be an incorrect representation of the target.

Incompleteness. A vehicle is an incomplete epistemic representation of a target if and only if, for some $o^{T}_{i} \in A^{T}$, o^{V}_{i} denotes o^{T}_{i} and $o^{V}_{i} \notin A^{V}$ or, for some ${}^{n}R^{T}_{k}$, o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$ and $<o_{1}^{T}, \ldots, o^{T}_{n}> \in {}^{n}R^{T}_{k}$. but $<o^{V}_{1}, \ldots, o_{n}^{V}> \notin {}^{n}R^{V}_{k}$. However, an intended partial isomorphism can hold between the relevant structure of the vehicle and that of the target even if, for some $o^{T}_{i} \in A^{T}$, o^{V}_{i} denotes o^{T}_{i} and $o^{V}_{i} \notin A^{V}$ or, for some ${}^{n}R^{T}_{k}$, o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R^{T}_{k}$ and $<o^{T}_{1}, \ldots, o_{k}^{T}> \in {}^{n}R_{k}^{T}$, but $<o^{V}_{1}, \ldots, o^{V}_{n}> \notin {}^{n}R^{V}_{k}$. For an intended partial isomorphism to hold between the relevant structure of the vehicle and that of the target, an object can be an element of the universe of the relevant structure of the target even if the objects that denotes it is not in the relevant structure of the vehicle—the first object will simply not be in the universe of the maximal intendedly isomorphic *partial sub*structure of the target. Analogously, for an intended partial isomorphism to hold between the relevant structure of the vehicle and that of the target, it is not necessary that, if o^{V}_{1} denotes o^{T}_{1}, …, o^{V}_{n} denotes o^{T}_{n}, ${}^{n}R^{V}_{k}$ denotes ${}^{n}R_{k}^{T}$ and $<o^{T}_{1}, \ldots, o^{T}_{n}> \in {}^{n}R_{k}^{T}$, $<o^{V}_{1}, \ldots, o^{V}_{n}> \in {}^{n}R^{V}_{k}$. If $<o^{T}_{1}, \ldots,$

$o^{T}{}_{n}> \in {}^{n}R_{k}^{T}$ and $<o^{V}{}_{1}, \ldots, o^{V}{}_{n}> \notin {}^{n}R^{V}{}_{k}$, it will simply be the case that $<o^{V^*}{}_{1}, \ldots, o^{V^*}{}_{n}> \in indet({}^{n}R^{V^*}{}_{k})$ and $<o^{T^*}{}_{1}, \ldots, o^{T^*}{}_{n}> \in indet({}^{n}R^{T^*}{}_{k})$. So, if an intended partial isomorphism holds between the relevant structure of the target and that of the vehicle, the vehicle can be an incomplete representation of the target.

III.2.12. Partial Isomorphism and Its Limit Cases

The partial isomorphism conception, not only seems to be able to deal with cases of incorrect and incomplete representation such as that of the old London Underground and today's network, but it also provides an explanation of why the other less successful proposals that we have considered so far succeed where they do. Isomorphism is a limit case of partial isomorphism in which the isomorphic partial substructures of the relevant structures of the vehicle and the target are the structures themselves (i.e. the universe of both structures are identical with the universe of their partial substructures and for all relations the set of n-tuples for which it is undetermined whether the n-tuples satisfy or do not satisfy the structure in question is empty).

The intended partial isomorphism account would thus be able to account for the success of the intended isomorphism conception as a conception of completely faithful epistemic representation. Intended isomorphism between the relevant structures of the vehicle and the target is a limit case of intended partial isomorphism between the two—it is the case in which the relevant structure of the vehicle and the target are themselves the maximal intendedly isomorphic partial substructures—analogously to the way in which completely faithful epistemic representation is the limit case of partially faithful epistemic representation—it is the case in which *all* inferences form the vehicle to the target are sound.

In analogous manner, the intended partial isomorphism conception allows us to explain why the intended homomorphism and intended counter-homomorphism conception are successful where they are successful. Intended homomorphisms and intended inverse homomorphisms are also limit cases of intended partial isomorphism. They are the cases in which the relevant "partial" substructures of the vehicle (in the case of intended homomorphism) and the target (in the case of intended inverse homomorphism) are identical to their relevant total structures.

This flexibility comes at a (small) price—one of the limit cases of partial isomorphism threatens to undermine the intended partial isomorphism conception of partially faithful epistemic representation and calls for the proviso I mentioned above. This is the case in which, for all relations and functions in the substructure and all n-tuples, it is undetermined whether the n-tuple satisfies or does not satisfy the relation (i.e. the case in which, for all R_i^{nP}, both $sat(R_i^{nP})=\varnothing$ and $dissat(R_i^{nP})=\varnothing$) and the value of the function for that n-tuple is undetermined. This case is a threat for the partial intended partial isomorphism conception because, if this limit case of partial isomorphism holds between the relevant structures of the vehicle and the target, then the user would not be able to draw from the vehicle any true conclusion about the target. If this limit case of intended partial isomorphism was allowed, completely unfaithful epistemic representation would count as partially faithful on the intended partial isomorphism account. To avoid becoming vacuous, the intended partial isomorphism account of faithful epistemic representation has to rule out this case as a case of partially faithful epistemic representation. This is the rationale behind the proviso that the intended partial isomorphism be non-vacuous.

(28) A partial isomorphism is *non-vacuous* if and only if:

[28.1] A^{V^*} is not empty and

[28.2]

[28.2.1] for some ${}^{n}R_i^{V^*}$ (if there are any), $sat({}^{n}R_i^{V^*})\neq\varnothing$ or $dissat({}^{n}R^{V^*}{}_i)\neq\varnothing$, or

[28.2.2] for some ${}^{n}F_i^{V^*}$ (if there are any), ${}^{n}F_i^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*})$ has a determinate value for some $o_1^{V^*}, \ldots, o_n^{V^*}$.

I will now argue that, if the intended partial isomorphism that holds between the relevant structure of the vehicle and that of the target is non-vacuous, then it must be the case that it is possible to perform some sound inferences from the vehicle to the target—and therefore the vehicle cannot be a completely unfaithful epistemic representation of the target. If an intended partial isomorphism holds between the vehicle and the target, then an intended isomorphism must hold between the maximal intendedly isomorphic partial substructure of the vehicle and that of the target. If the intended partial isomorphism is non-vacuous, then it must be the case that the universes of the maximal intendedly isomorphic partial substructures of the vehicle and the target are not empty. So, there must be at least one object in the universe of the maximal intendedly isomorphic partial substructure of the vehicle. Since an intended isomorphism holds between the two substructures, each of the objects in the universe of the maximal intendedly isomorphic partial substructure of the vehicle denotes one and only one object in the universe of the maximal intendedly isomorphic partial substructure of the target. From the definition of a non-vacuous intended partial isomorphism it follows that in the maximal intendedly isomorphic partial substructure of the vehicle there must be at least one *n*-ary relation that is determined for at least one *n*-tuple of objects or one *n*-ary function whose value is determined for at least one *n*-tuple of arguments.

I will now argue that if one of these two conditions holds, then there is at least one sound inference from the vehicle to the target and, therefore, the vehicle cannot be a completely unfaithful epistemic representation of the target. If there is one *n*-ary relation ${}^{n}R_k^{V^*}$ that is determined for one *n*-tuple of objects, $o_1^{V^*}, \ldots, o_n^{V^*}$, and an intended isomorphism holds between the maximal intendedly isomorphic partial substructure of the vehicle and that of the target, it follows that, in the maximal intendedly isomorphic partial substructure of the target, the objects denoted by, respectively, $o_1^{V^*}$, …, and $o_n^{V^*}$ are either in $sat({}^{n}R^{T^*}{}_k)$ or $dissat({}^{n}R^{T^*}{}_k)$, where ${}^{n}R^{V}{}_k$ denotes ${}^{n}R_k^{T}$. From the construction of the maximal intendedly isomorphic partial substructures of the vehicle and the target it follows that either both $<o_1^{V}, \ldots, o^{V}{}_n> \in {}^{n}R^{V}{}_k$ and $<o_1^{T}, \ldots, o_n^{T}> \in {}^{n}R_k^{T}$ (if $<o_1^{V^*}, \ldots, o_n^{V^*}> \in \mathrm{sat}({}^{n}R^{V^*}{}_k)$ and $<o_1^{T^*}, \ldots, o_n^{T^*}> \in sat({}^{n}R^{T^*}{}_k)$) or $<o_1^{V}, \ldots, o_n^{V}> \notin {}^{n}R_k^{V}$ and $<o^{T}{}_1, \ldots, o^{T}{}_n> \notin {}^{n}R_k^{T}$ (if $<o^{V^*}{}_1, \ldots, o_n^{V^*}> \in sat({}^{n}R^{V^*}{}_k)$ and $<o_1^{T^*}, \ldots, o_n^{T^*}> \in sat({}^{n}R_k^{T^*})$). In either case, there is at least one sound inference from the vehicle to the target—the inference from the fact that $o_1^{V}, \dots, o_k^{V}$ are (are not) in the relation ${}^{n}R_k^{V}$ to the fact that $o_1^{T}, \dots, o_k^{T}$ are (are not) in the relation ${}^{n}R_k^{T}$.

If there is an n-ary function whose value is determined for at least one n-tuple of arguments and an intended isomorphism holds between the maximal intendedly isomorphic partial substructure of the vehicle and that of the target, it follows that ${}^{n}F_k^{V^*}(o_1^{V^*}, \dots, o_n^{V^*})=o^{V^*}{}_i$ and ${}^{n}F_k^{T^*}(o_1^{T^*}, \dots, o_n^{T^*})=o^{T^*}{}_i$ for some n-tuple of objects, $o_1^{V^*}, \dots, o_n^{V^*}$. From the construction of the maximal intendedly isomorphic partial substructures of the vehicle and the target it follows that ${}^{n}F^{V}{}_k(o_1^{V}, \dots, o_n^{V})=o_i^{V}$ and ${}^{n}F^{T}{}_k(o^{T}{}_1, \dots, o^{T}{}_n)=o^{T}{}_i$. So, there is at least one sound inference from the vehicle to the target—the inference from the fact that the value of the function ${}^{n}F^{V}{}_k$ is $o^{V}{}_i$ for the arguments $o_1^{V}, \dots, o_n^{V}$ to the fact that that the value of the function ${}^{n}F^{T}{}_k$ is $o^{T}{}_i$ for the arguments $o^{T}{}_1, \dots, o_n^{T}$.

It follows that if the intended isomorphism between the maximal intendedly isomorphic partial substructures, V* and T*, of, respectively, the relevant structure of the vehicle, V, and that of the target, T, is non-vacuous, the vehicle cannot be a completely unfaithful epistemic representation of the target and therefore it must be to some degree a faithful epistemic representation of the target.

III.2.13. Partial Isomorphism and Inexactness

The intended partial isomorphism conception seems to be very successful in accounting for epistemic representations that are both incomplete and inaccurate. However, in this subsection, I will argue that the intended partial isomorphism conception is not equally successful in accounting for inexact representations. The case of inexact representations is particularly important for our purposes because inexactness seems to be a pervasive kind of unfaithfulness in models from mathematised sciences. We have seen this to be the case with the case of the inclined model and the soap-box derby. In that case, we had reasons to believe that the model was inexact in the sense that, for example, for some k and some t_i,

$f(g^{M}(b, t_i, k)) \neq g^{S}(f(b), f(t_i), f(k))$ (i.e. that some of the external forces on the box differed from the forces they denoted on the racer).

The problem, I think, is not that the intended partial isomorphism account of partially faithful epistemic representation cannot somehow accommodate inexact epistemic representations. Indeed, the intended partial isomorphism account has a strategy for dealing with inexact epistemic representations. The problem, I think, is that the strategy in question is not a good strategy.

From the definition of the maximal intendedly isomorphic partial substructures of the vehicle and the target, it follows that:

(i) ${}^{n}F_k^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*}) = o_i^{V^*}$ and ${}^{n}F_k^{T^*}(o_1^{T^*}, \ldots, o_n^{T^*}) = o_i^{T^*}$ if and only if both ${}^{n}F_k^{V}(o_1^{V}, \ldots, o_n^{V}) = o_i^{V}$ and ${}^{n}F_k^{T}(o_1^{T}, \ldots, o_n^{T}) = o_i^{T}$ and

(ii) ${}^{n}F_k^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*})$ and ${}^{n}F_k^{T^*}(o_1^{T^*}, \ldots, o_n^{T^*})$ are indeterminate if and only if ${}^{n}F_k^{V}(o_1^{V}, \ldots, o_n^{V}) \neq o_i^{V}$ or ${}^{n}F_k^{T}(o_1^{T}, \ldots, o_n^{T}) \neq o_i^{T}$.

If we adopt the convention that if ${}^{n}F^{V^*}{}_k(o^{V^*}{}_1, \ldots, o^{V^*}{}_n)$ and ${}^{n}F^{T^*}{}_k(o^{T^*}{}_1, \ldots, o^{T^*}{}_n)$ are both indeterminate, then $f({}^{n}F_k^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*})) = {}^{n}F_k^{T^*}(f(o_1^{V^*}), \ldots, f(o_n^{V^*}))$, then an intended partial isomorphism can hold among the relevant structures of the vehicle and the target even if the vehicle is an inexact representation of the target (where an epistemic representation is inexact if, for some ${}^{n}F^{V}{}_k$, $o^{V}{}_1$ denotes $o^{T}{}_1, \ldots, o^{V}{}_n$ denotes $o^{T}{}_n$, $o^{V}{}_i$ denotes $o^{T}{}_i$ and ${}^{n}F^{V}{}_k$ denotes ${}^{n}F^{T}{}_k$ and ${}^{n}F^{V}{}_k(o^{V}{}_1, \ldots, o^{V}{}_n) \neq o^{V}{}_i$ or ${}^{n}F^{T}{}_k(o^{T}{}_1, \ldots, o^{T}{}_n) \neq o^{T}{}_i$).

However, if this is the case, ${}^{n}F^{V^*}{}_k(o^{V^*}{}_1, \ldots, o^{V^*}{}_n)$ and ${}^{n}F^{T^*}{}_k(o^{T^*}{}_1, \ldots, o^{T^*}{}_n)$ will be indeterminate in the maximal intendedly isomorphic partial substructure of the vehicle and the target and, since we have assumed that if ${}^{n}F^{V^*}{}_k(o^{V^*}{}_1, \ldots, o^{V^*}{}_n)$ and ${}^{n}F^{T^*}{}_k(o^{T^*}{}_1, \ldots, o^{T^*}{}_n)$ are both indeterminate, then ${}^{n}F^{V^*}{}_k(o^{V^*}{}_1, \ldots, o^{V^*}{}_n) = {}^{n}F^{T^*}{}_k(o^{T^*}{}_1, \ldots, o^{T^*}{}_n)$, an intended isomorphism can hold between the maximal intendedly isomorphic partial substructure of the vehicle and that of the target.

It is important to note that this strategy treats inexactness as a case of incorrectness. An n-ary function ${}^{n}F^{V}{}_k(o^{V}{}_1, \ldots, o^{V}{}_n) = o^{V}{}_i$ can be construed as an $(n+1)$-ary relation $<o^{V}{}_1, \ldots, o^{V}{}_n, o^{V}{}_i> \in {}^{(n+1)}R^{V}{}_k$ and a partial function whose value is indeterminate for the arguments o^{V}_{1}, ..., o^{V}_{n} can be seen as a partial relation $<o^{V}_{1}, \ldots, o^{V}_{n}, o^{V}_{i}> \in undet(^{(n+1)}R^{V}_{k})$ for all o^{V}_{i}.

However, while treating inexactness as a kind of incorrectness may be a satisfactory strategy for dealing with some cases of inexactness, it is not the best strategy to deal with most of them. To see why in the case of scientific models, it is convenient to distinguish between approximations and idealizations. Some think that idealization is a kind of approximation. However, the rationale behind approximations and idealizations seems to be slightly different.

For example, the rationale behind not including in the inclined plane model a counterpart for forces such as the gravitational attraction of distant galaxies on the racer seems to be that their influence on the racer is negligible—their magnitude while different from zero is so close to zero that the effect of these forces on the racer is likely to be smaller than we are able to detect. The rationale behind not including in the model a counterpart for forces such as the air friction, on the other hand, is that these forces, even if not negligible, are very complicated to model and, often, the gain in faithfulness that derives from including a counterpart for them in the model is not worth the effort to do so.

Here I will call cases such as the first *approximations* and cases such as the second *idealizations*.[26] My contention is that the intended partial isomorphism strategy, which may be satisfactory in cases of idealization, is not equally satisfactory in cases of approximation. When it comes to approximations, it *does* matter that, even if $f(^{n}F^{V}_{k}(o^{V}_{1}, \ldots, o^{V}_{n}) \neq {}^{n}F^{T}_{k}(f(o^{V}_{1}), \ldots, f(o^{V}_{n}))$, $f(^{n}F^{V}_{k}(o^{V}_{1}, \ldots, o^{V}_{n})$ and ${}^{n}F^{T}_{k}(f(o^{V}_{1}), \ldots, f(o^{V}_{n}))$ are "close enough". For example, if the inclined plane model is to be a faithful epistemic representation of the soap-box derby, then it matters that the gravitational force exerted by every massive object in the universe on the racer is close to zero even if it is not identical to zero because in the model all the

[26] Obviously, I do not mean to offer an account of approximation or idealization; nor do I intend to offer a criterion to demarcate idealizations and approximations. I use 'approximation' and 'idealization' simply as two convenient labels to distinguish between two kinds of unfaithfulness.

forces on the box other than the gravitational force and the normal force exerted by the plane are set to zero.

More generally, the reason for distinguishing between incorrectness and inexactness is that the latter comes in degrees. A user can draw conclusions that are strictly speaking false from inexact and incorrect epistemic representations alike. However, inexact epistemic representations, unlike incorrect epistemic representations, may allow their users to draw conclusions that are "closer to the truth" or "farther from the truth" (in the sense mentioned in Section I.2.5 above).

So, for example, the position function of the box at a certain time and that of the racer at the corresponding time may differ more or less significantly and, everything else being equal, the less significant this difference is, the model would seem to be a more faithful epistemic representation of its target because the less significant the difference between the velocity function of the box at a certain time and that of the racer at the corresponding time, the more likely it is that a user would consider the inference from the velocity function of the box at a certain time to that of the racer at the corresponding time approximately true (and therefore is closer to the truth in the sense discussed in I.2.5).

On an account of partially faithful epistemic representation that considers inexactness a kind of incorrectness, however, all inexact representations would seem to be equally unfaithful. So, for example, no matter how small the difference between the difference between the velocity function of the box at a certain time and that of the racer at the corresponding time, the conclusion is equally incorrect. The problem of inexactness, which is so crucial in the case of models, can only be dealt with adequately by a conception that acknowledges the specific nature of inexact representations. In the next section, I intend to develop such a conception.

III.2.14. The Structural Similarity Account of Partially Faithful Epistemic Representation

In the rest of this chapter, I will develop a version of the structural conception of faithful epistemic representation that I call the *structural similarity account of* *faithful epistemic representation*. The structural similarity account of faithful epistemic representation maintains that

(M) a vehicle is a partially faithful, analytically interpreted epistemic representation of a certain target for a certain user if and only if:

(M.1) the vehicle is an analytically interpreted epistemic representation of the target for that user and

(M.2) there is a partial substructure of the relevant structure of the vehicle V and a partial substructure of the relevant structure of the target T such that an intended non-vacuous partial quasi-isomorphism holds between the relevant structure of the vehicle and that of the target.

A quasi-isomorphism is a morphism that is defined exactly like an isomorphism except for the fact that, for every ${}^{n}F^{A}_{k}$ and ${}^{n}F^{B}_{k}$, $f({}^{n}F^{A}_{k}(o_1{}^{A}, \ldots, o_n{}^{A}))$ does not need to be equal to ${}^{n}F^{B}_{k}(f(o_1{}^{A}), \ldots, f(o_n{}^{A}))$ for **A** to be quasi-isomorphic to **B**, they only need to be "close enough". Various forms of quasi-isomorphism are obtained from the definition of isomorphism (definition (16)) by substituting condition [16.4] with a weaker condition. Here I will develop *one* version of the notion of quasi-isomorphism. However, there may be other versions of quasi-isomorphism that suit the purposes of the structural similarity account equally well as the one I suggest here.

If $f({}^{n}F^{A}_{k}(o^{A}_{1}, \ldots, o^{A}_{n}))$ and ${}^{n}F^{B}_{k}(f(o^{A}_{1}), \ldots, f(o^{A}_{n}))$ are scalars or vectors, the *deviation* between $f({}^{n}F^{A}_{k}(o^{A}_{1}, \ldots, o^{A}_{n}))$ and ${}^{n}F^{B}_{k}(f(o^{A}_{1}), \ldots, f(o^{A}_{n}))$ is the non-negative real number d such that $\frac{(|{}^{n}F^{B}_{k}(f(o^{A}_{1}),\ldots,f(o^{A}_{n}))| - |f({}^{n}F^{A}_{k}(o^{A}_{1},\ldots,o^{A}_{n}))|)^2}{|{}^{n}F^{B}_{k}(f(o^{A}_{1}),\ldots,f(o^{A}_{n}))|} = d$. Here $|x|$ denotes the absolute value of x, if x is a scalar, and the non-negative norm of $|x|$, if x is a vector. (Note that by dividing the difference between ${}^{n}F^{B}_{i}(f(o^{A}_{1}), \ldots, f(o^{A}_{n})))$ and ${}^{n}F^{A}_{i}(o^{A}_{1}, \ldots, o^{A}_{n}))$ by ${}^{n}F^{B}_{i}(f(o^{A}_{1}), \ldots, f(o^{A}_{n})))$ we make the deviation between ${}^{n}F^{B}_{i}(f(o^{A}_{1}), \ldots,$

$f(o^A{}_n)))$ and $^nF^A{}_i(o^A{}_1, \ldots, o^A{}_n))$ independent from the magnitude of $^nF^B{}_i(f(o^A{}_1), \ldots, f(o^A{}_n)))$.

(29) A function, f, from A^A onto A^B is a *quasi-isomorphism of maximum deviation m* if and only if:

[29.1] For every $o_i^B \in A^B$, there is an $o_i^A \in A^A$ such that $f(o_i^A)=o_i^B$.

[29.2] For every $o_i^B \in A^B$, if $f(o_i^A)=o_i^B$ and $f(o_k^A)=o_i^B$, then $o_k^A=o_i^A$.

[29.3] For all R_i^A and R_i^B, $<o_1^A, \ldots, o_i^A> \in R_i^A$ if and only if $<f(o_1^A), \ldots, f(o_k^A)> \in R_i^B$.

[29.4] for all $^nR^A{}_k$ and $^nR^B{}_k$, $<o^A{}_1, \ldots, o^A{}_n> \in {}^nR^A{}_k$ if and only if $<f(o^A{}_1), \ldots, f(o^A{}_n)> \in {}^nR^B{}_k$.

[29.5] for all $^nF^A{}_k$ and $^nF^A{}_k$ and $^nF^B{}_1, \ldots, {}^nF^B{}_j$ such that $f(^nF^A{}_k(o^A{}_1, \ldots, o^A{}_n)$ and $^nF^B{}_k(f(o^A{}_1), \ldots, f(o^A{}_n))$ are either scalars or vectors,

$$\frac{(|^nF^B{}_k(f(o^A{}_1),\ldots,f(o^A{}_n))|-|f(^nF^A{}_k(o^A{}_1,\ldots,o^A{}_n))|)^2}{|^nF^B{}_k(f(o^A{}_1),\ldots,f(o^A{}_n))|} \leq m,$$

[29.6] $f(^nF^A{}_k(o^A{}_1, \ldots, o^A{}_n)$ is indeterminate if and only if $^nF^B{}_k(f(o^A{}_1), \ldots, f(o^A{}_n))$ is indeterminate.

[29.7] $f(^nF^A{}_k(o^A{}_1, \ldots, o^A{}_n)={}^nF^B{}_k(f(o^A{}_1), \ldots, f(o^A{}_n))$, otherwise.

A (non-vacuous) partial quasi-isomorphism holds between two structures if and only if a quasi-isomorphism holds between two (non-vacuous) partial substructures of those structures.

One of the crucial differences between the structural similarity account and the intended partial isomorphism account is that the intended morphism that holds between the partial substructures is an intended quasi-isomorphism rather than an intended isomorphism. Another crucial difference is that, on the structural similarity account, the maximal intendedly isomorphic partial substructures of the vehicle, V*, and that of the target, T*, are different from the ones put forward by the intended partial isomorphism account.

(30) If **V** and **T** are the relevant structure of the vehicle, *V*, and the target, *T*, relative to a certain interpretation, *I*, of the vehicle in terms of the target, **V*** and **T***, are the ***maximal intendedly quasi-isomorphic partial substructures*** of, respectively, **V** and **T** if and only if:

[30.1] **V*** is a partial substructure of **V**;

[30.2] **T*** is a partial substructure of **T**,

[30.3] If, according to *I*, o_i^V denotes o_i^T, $o^{V^*}{}_i \in A^{V^*}$ and $o_i^{T^*} \in A^{T^*}$ if and only if $o_i^V \in A^V$ and $o_i^T \in A^T$.

[30.4] If, according to *I*, o_1^V denotes o_1^T, ..., o_n^V denotes o_n^T, ${}^nR_k^V$ denotes ${}^nR_k^T$, $o_1^{V^*} \in A^{V^*}$, ..., $o_n^{V^*} \in A^{V^*}$, then:

[30.4.1] $<o_1^{V^*}, \ldots, o_n^{V^*}> \in sat({}^nR_k^{V^*})$ and $<o_1^{T^*}, \ldots, o_n^{T^*}> \in sat({}^nR_k^{T^*})$ if and only if $<o_1^V, \ldots, o_n^V> \in {}^nR_k^V$ and $<o_1^T, \ldots, o_n^T> \in {}^nR_k^T$,

[30.4.2] $<o_1^{V^*}, \ldots, o^{V^*}{}_n> \in dissat({}^nR^{V^*}{}_k)$ and $<o_1^{T^*}, \ldots, o_n^{T^*}> \in sat({}^nR_k^{T^*})$ if and only if $<o_1^V, \ldots, o_n^V> \notin {}^nR_k^V$ and $<o^T{}_1, \ldots, o^T{}_n> \notin {}^nR_k^T$, and

[30.4.3] $<o_1^{V^*}, \ldots, o_n^{V^*}> \in indet({}^nR_k^{V^*})$ and $<o_1^{V^*}, \ldots, o_n^{V^*}> \in dissat({}^nR_k^{V^*})$ if and only if either $<o_1^V, \ldots, o_n^V> \in {}^nR_k^V$ and $<o_1^T, \ldots, o_n^T> \notin {}^nR_k^T$ or $<o_1^V, \ldots, o_n^V> \notin {}^nR_k^V$ and $<o_1^T, \ldots, o_n^T> \in {}^nR_k^T$.

[30.5] If, according to *I*, o_1^V denotes o_1^T, ..., o_n^V denotes o_n^T, o_i^V denotes o_i^T, ${}^nF_k^V$ denotes ${}^nF_k^T$, $o_1^{V^*} \in A^{V^*}$, ..., $o_n^{V^*} \in A^{V^*}$, $o_1^{T^*} \in A^{T^*}$, ..., and $o_n^{T^*} \in A^{T^*}$, then:

[30.5.1] if $f({}^nF_k^V(o_1^V, \ldots, o_n^V))$ and ${}^nF_k^T(f(o_1^V), \ldots, f(o_n^V))$ are scalars or vectors ${}^nF_k^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*}) = o_i^{V^*}$ and ${}^nF_k^{T^*}(o_1^{T^*}, \ldots,$

$o_n^{T^*}) = o_i^{T^*}$ if and only if

$$\frac{(|{}^nF_k^{T}(f(o_1^{V}),...,f(o_n^{V}))| - |f({}^nF_k^{V}(o_1^{V},...,o_n^{V}))|)^2}{|{}^nF_k^{T}(f(o_1^{V}),...,f(o_n^{V}))|} \le D$$

[30.5.2] if $f({}^nF_k^V(o_1^V, \ldots, o_n^V))$ and ${}^nF_k^T(f(o_1^V), \ldots, f(o_n^V))$ are scalars or vectors, ${}^nF_k^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*})$ and ${}^nF_k^{T^*}(o_1^{T^*}, \ldots, o_n^{T^*})$ are indeterminate if and only if

$$\frac{(|{}^nF_k^{T}(f(o_1^{V}),...,f(o_n^{V}))| - |f({}^nF_k^{V}(o_1^{V},...,o_n^{V}))|)^2}{|{}^nF_k^{T}(f(o_1^{V}),...,f(o_n^{V}))|} > D.$$

[30.5.3] if $f({}^nF_k^V(o_1^V, \ldots, o_n^V))$ or ${}^nF_k^T(f(o_1^V), \ldots, f(o_n^V))$ is not a scalar or a vector, ${}^nF_k^{V^*}(o_1^{V^*}, \ldots, o_n^{V^*}) = o_i^{T^*}$ and ${}^nF_k^{T^*}(o_1^{T^*}, \ldots, o_n^{T^*}) = o_i^{T^*}$.

Condition [30.5] is what differentiates the definition of maximal intendedly quasi-isomorphic partial substructure (Definition (30)) from that of the partial isomorphism account's definition of maximal intendedly isomorphic partial substructure (Definition (27)). According to the structural similarity account, if $f({}^nF^V_k(o^V_1, \ldots, o^V_n))$ and ${}^nF^T_k(f(o^V_1), \ldots, f(o^V_n))$ are scalars or vectors, $f({}^nF^V_k(o^V_1, \ldots, o^V_n))$ and ${}^nF^T_k(f(o^V_1), \ldots, f(o^V_n))$ do not need to be identical in order for the corresponding functions in the maximal intendedly isomorphic partial substructures, ${}^nF^{V^*}_k(o^{V^*}_1, \ldots, o^{V^*}_n)$ and ${}^nF^{T^*}_k(o^{T^*}_1, \ldots, o^{T^*}_n)$, not to be indeterminate. The functions in the partial substructure, ${}^nF^{V^*}_k(o^{V^*}_1, \ldots, o^{V^*}_n)$ and ${}^nF^{T^*}_k(o^{T^*}_1, \ldots, o^{T^*}_n)$, are not indeterminate if the deviance between $f({}^nF^V_k(o^V_1, \ldots, o^V_n))$ and ${}^nF^T_k(f(o^V_1), \ldots, f(o^V_n))$ is less than or equal to D, which I take to be an ideal threshold between what we would consider an approximation and what we would consider an idealization.

So, according to the structural similarity account, for example, the inclined plane model is a partially faithful epistemic representation of the soap-box derby only if an intended non-vacuous partial quasi-isomorphism holds between the relevant structures of the inclined plane model and the soap-box derby. One of the intuitions that underlies the structural similarity account of faithful epistemic representation is that, everything else being equal, the closer to zero the average deviation of the quasi-isomorphism between the relevant substructure of the vehicle and that of the target, the more faithful a representation of the target the vehicle is.

Of course, a great deal more could be said here, incorporating insights from the literature on approximations—in particular with respect to the choice of the definition of divergence and the purpose-relativity of the size of *D*. But I hope that what I have developed here is sufficient to securing the chief philosophical point in the thesis and to lay a reasonable foundation for further work.

III.2.15. The Hierarchy of Morphisms, Structural Similarity and Degrees of Faithfulness

In this section, I will introduce the notion of structural similarity and clarify the relationship between the structural similarity account of faithful epistemic representation and the other versions of the structural conception of faithful epistemic representation that I have considered so far. First of all, it is important to note that two structures are partially isomorphic if and only if they are partially quasi-isomorphic with zero maximum deviance. The structural similarity conception can thus explain why the intended partial isomorphism conception is successful insofar as it is successful—partial isomorphism is a limit case of partial quasi-isomorphism for $m=0$. Since, as I have noted in Section III.2.12, isomorphism and homomorphism are themselves limit cases of partial isomorphism, they are also limit cases of partial quasi-isomorphism. This suggests that there is a hierarchy of morphisms of different strengths such that the stronger morphisms are limit cases of the weaker ones (see Fig. III.1).

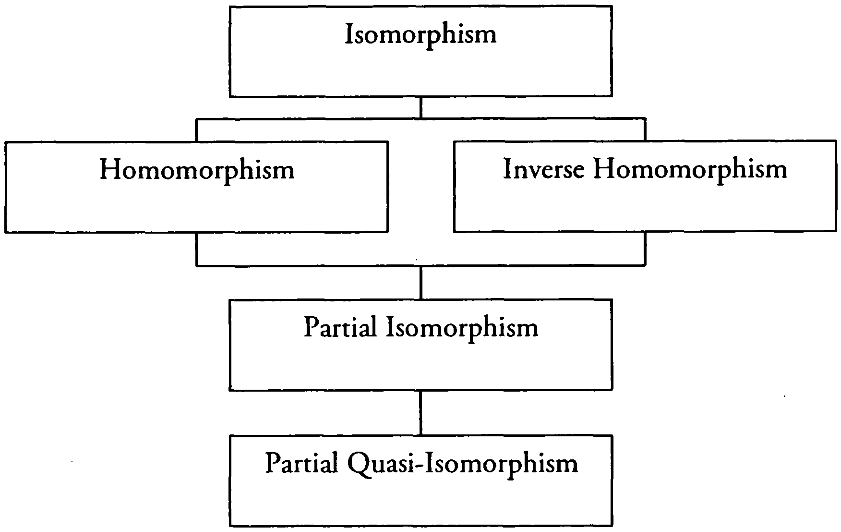


Figure III.1. The Hierarchy of Morphisms

We can thus introduce the notion of *(overall) structural similarity* between a vehicle and a target (with respect to a certain interpretation of the vehicle in terms of the target). Intuitively, the stronger the intended morphism between the relevant structure of the vehicle and that of the target, the more the vehicle and the target are structurally similar (with respect to a certain interpretation). Since isomorphism is the strongest morphism, the structural similarity between the vehicle and a target is *maximal* if an intended isomorphism holds between their relevant structures. The structural similarity is *minimal* in case the intended partial quasi-isomorphism meets only the one minimal set of requirements for being non-vacuous.

Between these two extremes, there is a spectrum of different degrees of structural similarity between a vehicle and a target.

(31) If A and B are two epistemic representations of the same target T, **A** is the relevant structure of A relative to interpretation I^A of A in terms of T, **B** is the relevant structure of B relative to interpretation I^B of B in terms of T, **TA** is the relevant structure of T relative to interpretations I^A of A in terms of T, **TB** is the relevant structure of T relative to interpretations I^B of B in terms of T, **A*** is

the maximal intendedly quasi-isomorphic partial substructure of **A**, **B*** is the maximal intendedly quasi-isomorphic partial substructure of **B**, and **TA*** and **TB*** are the maximal intendedly isomorphic partial substructures of, respectively, **TA** and **TB**, then A and B are *(overall) equally structurally similar* to T (relative to interpretations I^A and I^B) if and only if:

[31.1] the universe of **TB***, A^{TB^*}, is identical to the universe of **TB***, A^{TB^*},

[31.2] for all ${}^nR^{TA^*}{}_k$ and ${}^nR^{TB^*}{}_k$, $<o^{TA^*}{}_1, \ldots, o^{TA^*}{}_n>\in sat({}^nR^{TA^*}{}_k)$ if and only if $<o^{TB^*}{}_1, \ldots, o^{TB^*}{}_n>\in sat({}^nR^{TB^*}{}_k)$ and $<o^{TA^*}{}_1, \ldots, o^{TA^*}{}_n>\in dissat({}^nR^{TA^*}{}_k)$ if and only if $<o^{TB^*}{}_1, \ldots, o^{TB^*}{}_n>\in dissat({}^nR^{TB^*}{}_k)$, and

[31.3] for all ${}^nF^{TA^*}{}_k$ and ${}^nF^{TB^*}{}_k$, ${}^nF^{TA^*}{}_k(f(o^{A^*}{}_1)\ldots, f(o^{A^*}{}_n))={}^nF^{TB^*}{}_k(f(o^{B^*}{}_1)\ldots, f(o^{B^*}{}_n))$.

(32) If A and B are two epistemic representations of the same target T, **A** is the relevant structure of A relative to interpretation I^A of A in terms of T, **B** is the relevant structure of B relative to interpretation I^B of B in terms of T, **TA** is the relevant structure of T relative to interpretations I^A of A in terms of T, **TB** is the relevant structure of T relative to interpretations I^B of B in terms of T, **A*** is the maximal intendedly quasi-isomorphic partial substructure of **A**, **B*** is the maximal intendedly quasi-isomorphic partial substructure of **B**, and **TA*** and **TB*** are the maximal intendedly isomorphic partial substructures of, respectively **TA** and **TB**, then A is *(overall) more structurally similar* to T than B (relative to interpretations I^A and I^B) if:

[32.1] the universe of **TB***, A^{TB^*}, is a subset of the universe of **TB***, A^{TB^*}, [32.2] for all ${}^{n}R^{TA^*}{}_{k}$ and ${}^{n}R^{TB^*}{}_{k}$, if $<o^{TB^*}{}_{1}, \ldots, o^{TB^*}{}_{n}> \in sat({}^{n}R^{TB^*}{}_{k})$, then $<o^{TA^*}{}_{1}, \ldots, o^{TA^*}{}_{n}> \in sat({}^{n}R^{TA^*}{}_{k})$ and if $<o^{TB^*}{}_{1}, \ldots, o^{TB^*}{}_{n}> \in dissat({}^{n}R^{TB^*}{}_{k})$, then $<o^{TA^*}{}_{1}, \ldots, o^{TA^*}{}_{n}> \in dissat({}^{n}R^{TA^*}{}_{k})$,

[32.3] for all ${}^{n}F^{TA^*}{}_{k}$ and ${}^{n}F^{TB^*}{}_{k}$, if ${}^{n}F^{TA^*}{}_{k}(o^{TB^*}{}_{1}, \ldots, o^{TB^*}{}_{n})$ is determinate, then ${}^{n}F^{TA^*}{}_{k}(o^{TA^*}{}_{1}, \ldots, o^{TA^*}{}_{n})$ is determinate.

[32.4] for all ${}^{n}F^{TA^*}{}_{k}$ and ${}^{n}F^{TB^*}{}_{k}$ such that ${}^{n}F^{TA^*}{}_{k}(f(o^{A^*}{}_{1}), \ldots, f(o^{A^*}{}_{n}))$, $f({}^{n}F^{A^*}{}_{k}(o^{A^*}{}_{1}, \ldots, o^{A^*}{}_{n}))$, ${}^{n}F^{TB^*}{}_{k}(f(o^{B^*}{}_{1}), \ldots, f(o^{B^*}{}_{n}))$, and $f({}^{n}F^{B^*}{}_{k}(o^{B^*}{}_{1}, \ldots, o^{B^*}{}_{n}))$ are either scalars or vectors, then

$$\frac{(|{}^{n}F^{TA^*}{}_{k}(f(o^{A^*}{}_{1}),\ldots,f(o^{A^*}{}_{n}))| - |f({}^{n}F^{A^*}{}_{k}(o^{A^*}{}_{1},\ldots,o^{A^*}{}_{n}))|)^2}{|{}^{n}F^{TA^*}{}_{k}(f(o^{A^*}{}_{1}),\ldots,f(o^{A^*}{}_{n}))|} \leq$$

$$\frac{(|{}^{n}F^{TB^*}{}_{k}(f(o^{B^*}{}_{1}),\ldots,f(o^{B^*}{}_{n}))| - |f({}^{n}F^{B^*}{}_{k}(o^{B^*}{}_{1},\ldots,o^{A^*}{}_{n}))|)^2}{|{}^{n}F^{TB^*}{}_{k}(f(o^{B^*}{}_{1}),\ldots,f(o^{B^*}{}_{n}))|}$$

[32.5] and at least one of the following conditions is satisfied:

[32.5.1] A^{TB^*} is a proper subset of A^{TA^*},

[32.5.2] for some ${}^{n}R^{TA^*}{}_{k}$ and ${}^{n}R^{TB^*}{}_{k}$, $<o^{TA^*}{}_{1}, \ldots, o^{TA^*}{}_{n}> \in sat({}^{n}R^{TA^*}{}_{k})$ and $<o^{TB^*}{}_{1}, \ldots, o^{TB^*}{}_{n}> \in indet({}^{n}R^{TB^*}{}_{k})$ or $<o^{TA^*}{}_{1}, \ldots, o^{TA^*}{}_{n}> \in dissat({}^{n}R^{TA^*}{}_{k})$ and $<o^{TB^*}{}_{1}, \ldots, o^{TB^*}{}_{n}> \in indet({}^{n}R^{TB^*}{}_{k})$,

[32.5.3] for some ${}^{n}F^{TA^*}{}_{k}$ and ${}^{n}F^{TB^*}{}_{k}$, ${}^{n}F^{TA^*}{}_{k}(o^{TA^*}{}_{1}, \ldots, o^{TA^*}{}_{n})$ is determinate and ${}^{n}F^{TA^*}{}_{k}(o^{TB^*}{}_{1}, \ldots, o^{TB^*}{}_{n})$ is indeterminate,

[32.5.4] for some ${}^{n}F^{TA^*}{}_{k}$ and ${}^{n}F^{TB^*}{}_{k}$ such that ${}^{n}F^{TA^*}{}_{k}(f(o^{A^*}{}_{1}), \ldots, f(o^{A^*}{}_{n}))$, $f({}^{n}F^{A^*}{}_{k}(o^{A^*}{}_{1}, \ldots, o^{A^*}{}_{n}))$, ${}^{n}F^{TB^*}{}_{k}(f(o^{B^*}{}_{1}), \ldots, f(o^{B^*}{}_{n}))$, and $f({}^{n}F^{B^*}{}_{k}(o^{B^*}{}_{1}, \ldots, o^{B^*}{}_{n}))$ are either scalars or vectors,

$$\frac{(|{}^{n}F^{TA^*}{}_{k}(f(o^{A^*}{}_{1}),\ldots,f(o^{A^*}{}_{n}))| - |f({}^{n}F^{A^*}{}_{k}(o^{A^*}{}_{1},\ldots,o^{A^*}{}_{n}))|)^2}{|{}^{n}F^{TA^*}{}_{k}(f(o^{A^*}{}_{1}),\ldots,f(o^{A^*}{}_{n}))|}$$

$$< \frac{(|{}^{n}F^{TB^*}{}_{k}(f(o^{B^*}{}_{1}),\ldots,f(o^{B^*}{}_{n}))| - |f({}^{n}F^{B^*}{}_{k}(o^{B^*}{}_{1},\ldots,o^{B^*}{}_{n}))|)^2}{|{}^{n}F^{TB^*}{}_{k}(f(o^{B^*}{}_{1}),\ldots,f(o^{B^*}{}_{n}))|}$$

It is important to note that (32) is not a definition—it only gives us a set of sufficient conditions for a vehicle A to be more structurally similar to a target T

than a vehicle *B*. Whereas it is sufficient for a couple of vehicles to meet the set of conditions above in order for one of them to be a more structurally similar to a certain target than the other, it may well not be necessary to meet that set of conditions. There may well be other sets of conditions such that, if met by two vehicles, it would be reasonable to count the first as more structurally similar to the target than the second (with respect to certain interpretations of the two vehicles in terms of the target).

It is also important to note that the relation *being equally or more structurally similar to T than* (with respect to certain interpretations of the two vehicles in terms of the target) is a partial order—there are couples of representations of a certain target such that neither is more or equally structurally similar to the target than the other (under the interpretation of the respective vehicles in terms of the target).

The most fundamental intuition that underlies the structural similarity conception of faithful epistemic representation is that the more structurally similar a vehicle and a target are (with respect to a certain interpretation), the more faithful an epistemic representation of the target the vehicle is and vice versa—the more faithful a representation of the target the vehicle is, the more structurally similar a vehicle and a target are (with respect to a certain interpretation). More precisely, according to the structural similarity account of faithful epistemic representation:

(N) If *A* and *B* are two (analytically interpreted) epistemic representations of the same target *T*, *A* and *B* are equally faithful epistemic representations of *T* if and only if

(N.1) *A* and *B* are overall equally structurally similar to *T* (relative to interpretations I^A and I^B);

and

(O) If *A* and *B* are two (analytically interpreted) epistemic representations of the same target *T*, *A* is an overall more faithful epistemic representation of *T* (relative to interpretations I^A and I^B) than B (relative to interpretations)if and only if

(O.1) A is more structurally similar to T than B (relative to interpretations I^A and I^B).

On the structural similarity conception, for example, the new London Underground map is a more faithful epistemic representation of today's network than the old London Underground map by virtue of the fact that the new map is more structurally similar to today's network than the old map (with respect to the standard interpretation of the two maps in terms of the network). In other words, the new London Underground map is a more faithful epistemic representation of today's network than the old London Underground map because the intended morphism between the relevant structure of the new map and that of the network (i.e. intended isomorphism) is a stronger morphism than the one between the relevant structure of the old map and that of the network (i.e. intended partial isomorphism). In other words, in the case of the new map the maximal intendedly isomorphic partial structure of the relevant structure of the map and that of the relevant structure of network are identical to the very relevant structure of the map and that of the network. The same, however, is not true of the maximal intendedly isomorphic partial structure of the relevant structure of the map and that of the relevant structure of network.

What I want to emphasize here is that the idea that an intended partial quasi-isomorphism is a relation between two structures. The maximal intendedly quasi-isomorphic partial substructures are only supposed to highlight those portions of the relevant structures of the vehicle and of the target that are intendedly quasi-isomorphic so as to show how the structures of which they are partial substructures are structurally similar. So, the more the maximal intendedly quasi-isomorphic partial substructures of the relevant structure of the vehicle and that of the target are small portions of the relevant structures, the less we should think of the relevant structures as structurally similar.

Three remarks are in order here. First of all, the structural similarity account links structural similarity to the *overall* faithfulness of an epistemic representation to its target, which should be distinguished from what I have called the *specific* faithfulness of an epistemic representation to its target. Of two representations of a certain target, one can be overall more faithful than the other, if, say, one generates more (approximately) true conclusions than the other, and yet they can be equally faithful for some specific purpose, if from both we can draw some specific true conclusion which we happen to be interested in.

Second, it follows from the two theses above that, like the relation *being equally or more structurally similar to T than*, the relation *being (overall) an equally or more faithful epistemic representation of T than* is a partial order as well. That is, there may be pairs of representations of a certain target such that neither is equally or more faithful to the target than the other. This, however, is a desired result. As I have argued when introducing the notion of more and equally faithful epistemic representation in I.2.6, the relation *being (overall) an equally or more faithful epistemic representation of T than* is a partial order as well. That this should be the case is obvious when we consider that there are representations of a certain target that, by and large, represent different aspects of the same target. For example, on their standard interpretations, a view of Venice and a map of Venice both represent the city of Venice, but they represent largely different aspects of Venice. In a case like that, there seems to be no point in trying to determine which of two representations is more faithful to the target than the other.

Even if we only consider representations with largely the same scope, however, there are many cases in which we do not have any clear intuitions as how the unfaithfulness of the one compares to the unfaithfulness of the other. For example, we do not seem to have any clear intuition as to which of two partially faithful epistemic representations of a certain target is more faithful if one is a slightly incomplete but largely correct representation of the underground network while the other is a largely complete but to a certain extent incorrect representation of it. Is correctness weightier than completeness? And, if so, how much completeness is it worth to trade off for some correctness? Since there seems to be no uncontroversial or context or purpose neutral answer to these questions (at least outside of some specific context), it is possible to suggest that, in many cases, there is no way to compare the relative overall faithfulness of these two representations.

The fact that, on the structural similarity account, it is not always possible to determine which of two representations of a certain target is more faithful than the other is actually an advantage of the structural similarity account. In fact, I think that the best argument in favour of the direct connection between structural similarity and the overall faithfulness of a representation is that the structural similarity conception can successfully account for those cases in which we have clear intuitions about the relative overall faithfulness of two representations of a certain target, while it does not commit us to accepting that two representations are equally faithful or that one representation is more faithful that another in any cases in which our intuitions are not clear one way or the other. In the next section, for example, I will illustrate how the structural similarity conception can account for the fact that an inclined plane model with air friction is a more faithful epistemic representation of the soap-box derby system, than the inclined plane model we have considered so far on the grounds that the former is more structurally similar to the soap-box derby than the latter.

But, what if there were cases in which our intuitions about the overall faithfulness of two specific representations are uncontroversial that are not covered by the conditions above? As an advocate of the structural similarity conception, I would be committed to claiming that these can still be accounted for in terms of the relative structural similarity of two representations. This could be done by adding further sets of jointly sufficient conditions for a representation to be more structurally similar than another to a certain target overall or, on a case-by-case basis, by showing that the more faithful epistemic representation is overall more structurally similar to the target than the less faithful one. Consider for example the case of two maps of the London Underground one of which is correct and complete except for the fact that there is no circle or tab corresponding to one station, say, Piccadilly Circus, the other contains a circle or a tab for every station on the network but many of the properties of and relations among circles and tabs do not have a counterpart in the corresponding properties of and relations among the corresponding stations. Intuitively, I think we would tend to think that, even if neither map is a completely faithful epistemic representation of the London Underground network, the first map is a much more faithful epistemic representation of the network than the second one. The first map however does not meet condition ([32.1]) above to be more structurally similar to the network than the second map. However, as I have already noted, whereas it is sufficient that two representations meet those conditions in order for one to be more structurally similar to the target than the other, it is not necessary to meet that specific set of conditions for one to be more structurally similar than the other. Once one gets an intuitive grasp of the notion of structural similarity, they are likely to concede that the widespread incorrectness of the second map outweighs the slight incompleteness of the first and, therefore, the first map is more structurally similar to the network than the second even if neither map meets the set of conditions to be more structurally similar to the network than the other.

III.2.16. The Inclined Plane Model With and Without Air Friction

In this section, I illustrate how the structural similarity conception can account for our intuitions that a certain model is a more faithful epistemic representation of a certain system than another. So far we have considered the inclined plane model as a partially faithful epistemic representation of the soap-box derby system. In the model, there are only two forces acting on the box. The first is the gravitational force and the second is the normal force that the plane exerts on the box. In the system, on the other hand there are many more forces acting on the racer than the gravitational pull of the Earth and the normal force that the road exerts on the wheels of the racer. This is one of the reasons why the inclined plane model is only a highly idealized and abstract representation of the soap-box derby system.

A less idealized and more faithful epistemic representation of the soap-box derby, for example, could be obtained by using a model in which a third force acts on the box—the force due to air friction. Air friction is a difficult force to model accurately as the force due to air friction on a certain object depends on a large number of factors, including the velocity of the object, its shape, and the density of air. In the inclined plane model, we set the force on the box due to air friction at t_i equal to $-1/2C\rho A\ v^{M\wedge}(b, t_i)^2$, where A is the cross-sectional area of the box, ρ the air density, and C a dimensionless constant. The structure-type of the inclined plane model with air friction, M^, is thus analogous to the one of the inclined plane model without air friction except for the fact that in the relevant structure of the inclined plane model with friction $g^{M\wedge}(b, t_i, 3)=-1/2C\rho A\ v^{M\wedge}(b, t_i)^2)$.

Intuitively, we would tend to think that the inclined plane model with air friction is *overall* a more faithful epistemic representation of the soap-box derby then the inclined plane model without air friction, for the former takes into account one of the main factors that govern the behaviour of the racers, while taking also into account all other factors taken into account by the latter model. The structural similarity account of faithful epistemic representation vindicates the intuition that the inclined plane model with friction is a more faithful epistemic representation of the soap-box derby than the inclined plane model without friction. According to the structural similarity account this is due to the fact that the former is more structurally similar to the soap-box derby than the latter.

The inclined plane model with air friction in fact trivially meets conditions [32.1] and [32.2] above because, respectively, the relevant structures of the two models contain the same objects and do not contain any properties or relations. The functions $g^M(b, t_i, 1)$ and $g^{M\wedge}(b, t_i, 2)$, and $g^M(b, t_i, 2)$ and $g^{M\wedge}(b, t_i, 2)$ trivially meet condition [32.3] because the gravitational and normal forces acting on the box in the two models are the same and therefore their deviation from the corresponding forces on the racers are the same. The divergence between the air friction on the box at t_i (the function $g^{M\wedge}(b, t_i, 3)$) and the air friction on the racer at the corresponding time (the function $(g^S(f(b), f(t_i), f(3)))$ is less than the one between the function $g^M(b, t_i, 3)$ and $g^S(f(b), f(t_i), f(3))$ (because we have assumed that $g^M(b, t_i, 3)$ is set to 0 like all forces that act on the box in the inclined plane model without friction other than the gravitational and the normal force). As a result, the divergence between position of the box at t_i in the inclined plane model with air friction (function $\mathbf{r}^{M^\wedge}(b, t_i)$ and the average position of one of the racers at the corresponding time ($\mathbf{r}^S(b, t_i)$) is also likely to be smaller than the one between the position of the box in the inclined plane model without air friction at the same time and that of one of the racers. So, the inclined plane model with air friction is likely to be more structurally similar to the soap-box derby than the inclined plane model without friction in virtue of meeting conditions [32.1], [32.2], [32.3], [32.4]and [32.5].

Conclusions

In this thesis, I have identified a number of problems that are usually conflated under the heading of "the problem of scientific representation" and I have focussed on two of these problems, which, I have argued, are instances of two more general problems—i.e. the problem of what makes a certain vehicle an epistemic representation of a certain target and the problem of what makes a certain epistemic representation of a certain target a faithful epistemic representation of it.

Most of this thesis was devoted to laying the foundations for a solution for each of these problems. According to the account of epistemic representation that I have defended, the interpretational account of epistemic representation, a vehicle is an epistemic representation of a certain target for a certain user if and only if the user takes the vehicle to denote the target and she adopts an interpretation of vehicle (in terms of the target). In this thesis, I have focussed exclusively on one specific kind of interpretation, which I have called analytic interpretations. Whether the account I have developed can be developed into a general account of epistemic representation crucially depends on whether every possible interpretation can be reconstructed as an analytic interpretation. This is one of the crucial issues that are left open by this thesis and on which further work needs to be done.

One of the main advantages of the interpretational account is that it sheds light on the relation between epistemic representation and valid surrogative reasoning. As I have argued, the fact that a user adopts an interpretation of the vehicle in terms of the target (and takes the vehicle to stand for the target) is both that in virtue of what the vehicle is an epistemic representation of the target for her and that in virtue of what she can perform valid inferences from the vehicle to the target. Without the notion of an interpretation (or some analogous notion), the intimate relation between epistemic representation and valid surrogative reasoning remains unnecessarily mysterious.

The notion of an analytic interpretation plays also a central role in the account of faithful epistemic representation as it directly contributes to the solution to two crucial problems that have haunted the structuralist conception of (faithful) epistemic representation and indirectly to the solution of a third problem. The first problem is that of applying the notion of a morphism to objects that are not set-theoretic structure. As I have argued, the notion of an analytic interpretation provides us with a principled way to reconstruct the vehicle and the target as set-theoretic structures, which I have called respectively the relevant structure of the vehicle and the relevant structure of the target.

The second problem is that of determining which morphisms need to obtain between the relevant structure of the vehicle and that of the target in order for the vehicle to be a faithful epistemic representation of the target (to a certain degree) because a morphism may obtain between the relevant structure of the vehicle and that of the target without the first being a faithful epistemic representation of the target (on a certain interpretation of it). The notion of analytic interpretation provides us with a principled way to single out some of the morphism that may obtain between the relevant structure of the vehicle and that of the target as the intended morphism—i.e. the only ones that are relevant to the faithfulness of the epistemic representation in question.

The third problem is that the notion of the faithfulness of an epistemic representation comes in degrees while two structures are either *X*-morphic or they are not. However, the account of faithful epistemic representation that I have developed uses the notions of relvant structure and intended morphism to develop a third crucial notion that of the structural similarity between the vehicle and the target (under a certain interpretation of the former in terms of the latter). Intuitively, the stronger the strongest intended morphism between the relevant structure of the vehicle and that of the target is, the more structurally similar the vehicle and the target are. The central idea that underlies the structural similarity account is that the more structurally similar the vehicle and the target are under a certain interpretation of the former in terms of the latter, the more faithful an epistemic representation of the latter the former is under that interpretation. The account of faithful epistemic representation that I have developed, I think, vindicates the intuitions that underlie two of the main conceptions of representation but avoids the pitfalls that characterize the other versions of these views that I have considered.

The account of epistemic representation and that of faithful epistemic representation that I have developed and defended in this thesis are more than just compelementary—they are deeply interconnected. It is only when one attempts to develop some of the intuitions and ideas that can be found in the literature into a coherent whole that one can see how everything falls into place in the overall picture. In this thesis, I hope to have provided a good initial sketch of that picture.

Bibliography

Andrade, Edward Neville da Costa (1965). *Rutherford and the Nature of the Atom*, London: Heinman.

Bailer-Jones, Daniela M. (1999). 'Tracing the Development of Models in the Philosophy of Science', in L. Magnani, N. J. Nersessian und P. Thagard (Eds.), *Model-Based Reasoning in Scientific Discovery*, New York: Kluwer Academic/Plenum Publishers, 23–40.

Bailer-Jones, Daniela M. (2000). 'Modelling Extended Extragalactic Radio Sources', *Studies in History and Philosophy of Modern Physics*, 31B, S. 49-74.

Bailer-Jones, Daniela M. (2002). 'Scientists' Thoughts on Scientific Models', *Perspectives on Science* 10, 275–301.

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

Balzer, Wolfgang, Moulines, Ulises C., and Sneed, Joseph D. (1987). *An Architectonic for Science: The Structuralist Program*, Dordrecht: Reidel.

Bohr, Niels (1913). 'On the Constitution of Atoms and Molecules', *Philosophical Magazine*, Series 6, 26, 1–25.

Brown, James R. (1991). *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*, London: Routledge.

Callender, Craig and Jonathan Cohen (2006). 'There is No Problem of Scientific Representation', *Theoria* 55, 67–85.

Cartwright, Nancy (1983). *How The Laws of Physics Lie*, Oxford: Clarendon Press.

Cartwright, Nancy (1999a). *The Dappled World: A Study of the Limits of Science*, Cambridge: Cambridge University Press.

Cartwright, Nancy (1999b). 'Models and the Limits of Theory: Quantum Hamiltonians and the BCS Model of Superconductivity', in Morgan, M.S. and M. Morrison (eds.) (1999). *Models as Mediators: Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press, 241–281.

Cartwright, Nancy, Twofic Shomar and Mauricio Suárez (1995). 'The Tool-Box of Science: Tools for the Building of Models with a Superconductivity Example', *Poznan Studies in the Philosophy of the Sciences and the Humanities* 44, 137–149.

Chakravartty, Anjan (2001): 'The Semantic or Model-Theoretic View of Theories and Scientific Realism', *Synthese* 127, 325–345.

da Costa, Newton and French, Steven (2003). *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning,* Oxford: Oxford University Press.

da Costa, Newton and Steven French (1990). 'The Model-Theoretic Approach in the Philosophy of Science', Philosophy of Science 57, 248–265.

da Costa, Newton and Steven French (2000). 'Models, Theories and Structures: Thirty Years On', *Philosophy of Science* (Supplement) 67, S116–S127.

Fisher, Irving and Miller, O. M. (1944). *World Maps and Globes,* New York: Essential Books.

Fowler, Michael (1997). 'Rutherford Experiment and the Beginning of Nuclear Physics', Modern Physics 252 Lecture Notes, Charlottesville: University of Virignia.

French, Steve and Juha Saatsi (2004). 'Realism About Structure: The Semantic View and Non-Linguistic Representations' in *Proceedings of the Philosophy of Science Association 19th Biennial Meeting – PSA 2004: Contributed Papers,* Austin, TX.

French, Steven (2000). 'The Reasonable Effectiveness of Mathematics: Partial Structures and the Application of Group Theory to Physics', *Synthese* 125, 103–120.

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.

Friedman, Michael (1974). 'Explanation and Scientific Understanding', *The Journal of Philosophy,* 71, 5–19.

Frigg, Roman (2002). 'Models and Representation: Why Structures are not enough' *Measurement in Physics and Economics Discussion Paper Series,* DP MEAS 25/02, London: Centre for the Philosophy of Natural and Social Sciences.

Frigg, Roman (2003). *Re-presenting Scientific Representation,* Ph.D. Dissertation, London: University of London.

Frigg, Roman (2006). 'Scientific Representation and the Semantic View of Theories', *Theoria*, 55, 49–65.

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

Gendler, Tamar Szabó and Hawthorne, John (2002). *Conceivability and Possibility*. Oxford, Oxford University Press.

Giere, Ronald N. (1985). 'Constructive Realism' in P.M. Churchland and C. Hooker (eds.), *Images of Science. Essays on Realism and Empiricism with a Reply from Bas C. van Fraassen*. Chicago: University of Chicago Press, 75–98.

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

Giere, Ronald N. (1997). *Understanding Scientific Reasoning*, 4th edition, Fort Worth: Harcourt Brace College Publishers.

Giere, Ronald N. (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.

Giere, Ronald N. (2004). 'How Models are Used to Represent Reality', *Philosophy of Science*, 71, 742–752.

Goodman, Nelson (1968). *The Languages of Art: An Approach to a Theory of Symbols*, Indianapolis: Hackett.

Goodman, Nelson (1970). 'Seven Strictures on Similarity', in L. Foster and J.W. Swanson (eds.), *Experience and Theory*, London: Duckworth 19–29.

Hughes, R.I.G. (1997). 'Models and Representation', PSA 1996: Proceedings of the 1996 Biennial Meeting of the Philosophy of Science Association, 2, S325–S336.

Hughes, R.I.G. (1999). 'The Ising Model, Computer Simulation, and Universal Physics' in M. S. Morgan and M. Morrison (eds.), *Models as Mediators: Perspectives on Natural and Social Science*, Cambridge, Cambridge University Press, 97–145.

Knuuttila, T. (2004). 'Models, Representation, and Mediation', in *Proceedings of the 19th Biennial Meeting of the Philosophy of Science Association*.

Lewis, David (1969). *Convention: A Philosophical Study*, Cambridge, MA.: Harvard University Press.

Lopes, Dominic (1996). *Understanding Pictures*, Oxford: Oxford University Press.

Maling, D. H. (1992). *Coordinate Systems and Map Projections* (2nd edition), Oxford: Pergamon Press.

McMullin, Ernan (1985). 'Galilean Idealization', *Studies in History and Philosophy of Science*, 16, 247–273.

Morgan, Mary S. (2001), 'Models, Stories and the Economic World', *Journal of Economic Methodology*, 8, 361-384.

Morgan, Mary S. and Margaret Morrison, eds. (1999). *Models as Mediators: Perspectives on Natural and Social Science*, Cambridge, Cambridge University Press.

Morrison, Margaret (1997). 'Modelling Nature: Between Physics and the Physical World', *Philosophia Naturalis*, 35, 65–85.

Morrison, Margaret (1999). 'Models as Autonomous Agents', in M.S. Morgan and M. Morrison, eds. (1999). *Models as Mediators: Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press, 38–65.

Morrison, Margaret (forthcoming). 'Models as Representational Structures', in L. Bovens and S. Hartmann (eds.) *Nancy Cartwright's Philosophy of Science.*

Monmonier, Mark S. (1977). *Maps, Distortion, and Meaning*, Washington: Association of American Geographers.

Petruccioli, Sandro (1993). *Atoms, Metaphors, and Paradoxes: Niels Bohr and the Construction of a New Physics*, Cambridge: Cambridge University Press.

Rutherford, Ernest (1911). 'The Scattering of α and β Particles by Matter and the Structure of the Atom', *Philosophical Magazine*, Series 6, 21, 669–688.

Rutherford, Ernest (1914). 'The Structure of the Atom', *Philosophical Magazine (Series 6)*, 27, 488–498.

Sismondo, Sergio (1999). 'Models, Simulations, and their Objects', *Science in Context*, 12, 247–260.

Sismondo, Sergio and Nicholas Chrisman (2001) 'Deflationary Metaphysics and the Natures of Maps', *Philosophy of Science*, 68 (Proceedings), S38–S49.

Suárez, Mauricio (2002). 'The Pragmatics of Scientific Representation', CPNSS *Discussion Paper Series*, DP 66/02.

Suárez, Mauricio (2003). 'Scientific Representation: Similarity and Isomorphism' *International Studies in the Philosophy of Science*, 17, 225–244.

Suárez, Mauricio (2004). 'An Inferential Conception of Scientific Representation' *Philosophy of Science*, 71, 767–779.

Suppe, Frederick, ed. (1974a) *The Structure of Scientific Theories*, Urbana, IL.: University of Illinois Press.

Suppe, Frederick (1974b). 'The Search for Philosophic Understanding of Scientific Theories' in (Suppes 1974a), 3–233.

Suppe, Frederick (1989). *The Semantic Conception of Theories and Scientific Realism*, Urbana, IL.: University of Illinois Press.

Suppes, Patrick (1960). 'A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences', *Synthese*, 12, 287–301.

Suppes, Patrick (2002). *Representation and Invariance of Scientific Structures*, Stanford, CA.: CSLI Publications.

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.

van Dalen, Dirk (1994). *Logic and Structure*, 3rd edition, Berlin: Springer Verlag.

van Fraassen, Bas C. (1980). *The Scientific Image*, Oxford: Oxford University Press.

van Fraassen, Bas C. (1985) "Empiricism in the Philosophy of Science" in (Churchland and Hooker 1985), 245–308.

van Fraassen, Bas C. (1989). *Laws and Symmetry*, Oxford: Oxford University Press.

van Fraassen, Bas C. (1997) "Structure and Perspective: Philosophical Perplexity and Paradox", in M. L. Dalla Chiara et al. (eds.) *Logic and Scientific Methods*, Dordrecht: Kluwer, pp. 511–530.