How do models give us knowledge? The case of Carnot's ideal heat engine

TARJA KNUUTTILA AND MIEKE BOON

Department of Philosophy, University of Helsinki, Helsinki, Finland

Department of Philosophy, University of Twente, Enschede, The Netherlands

Abstract:

Our concern is in explaining how and why models give us useful knowledge. We argue that if we are to understand *how models function in the actual scientific practice* the representational approach to models proves either misleading or too minimal – depending on how representation is defined. With representational approach we refer to those approaches that attribute the epistemic value of models to the representational relationship between a model and some real target system. In contrast we propose turning from the representational approach to the artefactual, which implies also a new unit of analysis: the activity of modelling. Modelling, we suggest, could fruitfully be approached as a scientific practice in which concrete artefacts, i.e., models, are constructed with specific representational means and used in various ways, for example, for the purposes of scientific reasoning, theory construction and design of experiments and other artefacts. Furthermore, we propose that in the activity of modelling the model construction is intertwined with the construction of new phenomena, theoretical principles and new scientific concepts. We will illustrate these claims by studying the construction of the ideal heat engine by Sadi Carnot.

1 . Introduction

If there is any theme that unites philosophers as regards models, it is that of representation: Models are generally taken as representations. While scientific models as specifically designed artefacts certainly belong to the class of public objects called representations, something more is implied by the idea of representation in the context of modelling. Namely, the claim that models are representations plays out their relational nature. Models are typically conceived of as *models of* some natural phenomena and entities, and their epistemic value is assumed to be due to this representational relationship between them and their "target systems". To be sure, models may represent also theories or data (as data models). Yet the recent discussion on models and representation has concentrated on their relationship to real world systems, motivated by the general agreement among philosophers of science that models give us knowledge by virtue of representing (some selected aspects of) external world sufficiently accurately (e.g., Bailer-Jones 2003; da Costa and French 2000; French and Ladyman 1999; Frigg 2002; Morrison and Morgan 1999; Suárez 1999; Giere 2004; Contessa 2007, Mäki 2009). This agreement disguises, however, the fact that different philosophers understand representation in very different ways: While some philosophers have sought to ground the representational relationship in the respective properties of models and their target systems, others have claimed that representation is essentially an accomplishment of representation users. Yet, as we will argue in the following, neither approach really succeeds in accounting for the epistemic value of models: The former approach imposes too strong requirements as regards actual scientific representations, whereas the latter approach in turn ends up in presenting an overly minimalist account of representation.

To offer an alternative way of approaching models, better suited to tackle their function in knowledge-acquisition we consider models as epistemic tools. The "as… " locution serves to point out that we are not primarily interested in defining what models are – or are not. Approaching models "as" epistemic tools we rather wish to invoke the dimension of their use. Neither do we wish to dispute the fact that models often are *used to* represent some real target systems. Rather, we suggest that scientific models could be usefully approached through their artefactual dimension as constructed entities that give theoretical interpretations of some target phenomena in view of particular epistemic purposes. From this perspective it becomes more comprehensible, we suggest, how we gain knowledge by way of constructing models. The representational approach offers few resources to tackle the epistemic value of model construction as it focuses on the relationship between a ready-made model and its target system. Neither does it pay attention to how the target phenomena is typically (co-)constructed along the model, or how model construction often involves conceptual innovations.[i]

Our suggested turn from the representational approach to the artefactual implies also a new unit of analysis: the activity of modelling. Modelling, we suggest, could be seen as a scientific practice in which concrete artefacts, i.e., models, are constructed with the help of specific representational means and used in various ways, for example, for the purposes of scientific reasoning, theory construction and design of experiments and other artefacts.[ii] Furthermore, in this activity of modelling the model construction is intertwined with the construction of new phenomena, theoretical principles and scientific concepts. Thus justification of a model is partly built into it in the process of modelling, implying that the representational approach, albeit its focus on justification, fails to notice a lot of how models are justified in scientific practice.

In order to give a practical example of how our account of models as epistemic tools enhances our understanding of modelling, we study the construction of the Carnot model of a heat engine, which is a classical example of a technological device that was subject to scientific modelling. Although our account of Carnot's modelling is based on Carnot's writings, we do not present it as an accurate historical description of how Carnot actually went on in his modelling task but wish to remind our readers that it is inevitably a philosophical reconstruction. What is more, Carnot himself did not call his theoretical account of the heat-engine a model. The notion of a scientific model in its present sense was not in use those days (cf. Bailer-Jones, 1999). Thus it is only with the benefit of hindsight that the scientific community calls it a model of the heat-engine.

Our choice of the Carnot model is in part due to its familiarity, historical import and simplicity, but in part we have selected it for strategic reasons. Often the real target systems that models are supposed to represent are not known by us, but in the case of a heat engine a definite real target system exists. Furthermore, since it is a thing constructed by us, we should be able to assess in which sense the Carnot model of a heat engine can be considered an accurate, although selective, representation of the real heat engine. A possible objection to this case might be that being an engineering model the Carnot model has not the same status as scientific models in theoretical sciences proper. We think that such an objection is not valid for the reason that engineering sciences should not be confused with engineering.

Whereas the former concerns scientific research in the context of technological applications - which should become clear as our example unfolds - the latter is engaged in concrete design and development.

2. Models as Representations

According to the received wisdom models give us knowledge because they represent their supposed external target objects more or less accurately, in relevant respects and to a sufficient degree. This kind of formulation already suggests that there is a special sort of relationship between a model and its target, which is most commonly analysed in terms of similarity or isomorphism. Should the analysis of representation be based on such notions only, or is something else needed to establish the representational relationship? To this question philosophers of science have given various answers, which have far-reaching implications for how the epistemic value of models is to be understood.

The conviction that representation can be accounted for by reverting solely to the properties of the model and its target system, is part and parcel of the semantic approach to scientific modelling. According to the semantic conception, models specify structures that are posited as possible representations of either the observable phenomena or, even more ambitiously, the underlying structures of real target systems. The representational relationship between a model and its target system is analysed in terms of morphisms of various kinds: The suggested morphisms include isomorphism (e.g. van Fraassen 1980, Suppe 1974), partial isomorphism (e.g. da Costa and French 2000, French and Ladyman 1999) and homomorphism (Bartels 2006, Ambrosio 2007). The structuralist conception of scientific representation is usually cast in terms of isomorphism: a given structure represents its target system if they are structurally isomorphic to each other. Giere (1988) suggested similarity as an analysis of representational relationship in his reformulation of the semantic approach, but he has later come to think that similarity fits better the pragmatic approach to scientific representation (see below).

The semantic conception of models as a putative analysis of representation gives rise to persistent philosophical problems. Firstly, isomorphism does not have the right formal properties to capture the nature of the representational relationship: it is a symmetric, transitive and reflexive relationship whereas representation is not. Secondly, it does not leave room for misrepresentation. The idea that representation is either an accurate depiction of its object or not a representation at all does not fit actual representational practice. Thirdly, structure sharing is not necessary for representation. Scientific practice is full of examples of inaccurate models, which are difficult to render as isomorphic, or even partially isomorphic, to their targets. Fourthly, and perhaps most importantly, isomorphism does not capture the directionality of representation. (For extended discussion on these points see Suarez 1999, 2003, 2010; Frigg 2002, 2006).

Many of these problems are directly related to the fact that scientific representation is a relation between a representational vehicle (e.g., a model) and a real target, and thus a mere mathematical relation between two structures fails to capture some of its inherent features – and makes to too stringent demands on actual scientific representations. According to the pragmatists these problems will be cured if it is recognized that representation cannot be based only on the respective properties of the representational vehicle and its target system. What makes one thing a representation of another, i.e. what establishes the representational relationship, is the intended use by representation users. Thus representation becomes less a feature of models and their target systems than an accomplishment of representation users (Suárez 2004, 2010, Giere 2004, 2010, Bailer-Jones 2003, 2009).

As pragmatic approaches ground representation on the specific purposes and representing activity of humans, they cannot say anything substantive about the relationship of representation in general. This has also been explicitly admitted by the proponents of the pragmatic approach (see Giere 2004, Suárez 2004, van Fraassen 2008), of whom Suárez has gone farthest in arguing for a deflationist inferential account of representation which resists saying anything substantive about the supposed basis on which the representational power of representative vehicles rests, i.e. whether it rests, for instance, on isomorphism, similarity or denotation (for an account of scientific representation that builds on the notion of denotation, see Hughes 1999). Consequently, if we accept the deflationist pragmatist approach to representation, not much is established in claiming that models give us knowledge *because* they represent their target objects. While it appears to us that the pragmatist account offers most that can be said about representation at a general philosophical level, it makes the representational approach too minimal as an explanation of how we can gain knowledge through models. [iii]

In sum, as regards the representational approach to models we face the following dilemma: Either we attempt to offer representation a substantive account in terms of the properties of the model and its target that is able to explain in virtue of what do models give us knowledge – but fails for the reasons mentioned above. Or, alternatively, we settle for the pragmatist deflationist accounts, which stop short of being informative as to the epistemic value of models. Moreover, in their focus on representational relation between a ready-made model and its real target system neither of these accounts seems too fruitful in understanding the knowledge-bearing nature of *modelling*. The representational approach neither makes room for the study of how models are justified through their construction.

3. Models as Epistemic Tools

Recently, it has been questioned whether the link between models and their supposed target systems is as straightforward as the representational approach to models seems to imply. Apart from simplifications, approximations and idealizations, scientific modelling involves significant conceptual element, which covers such epistemic activities as discerning specific types of phenomena, conceptualizing 'non-directly observable' objects, properties, or processes, and bringing phenomena under specific types of 'non-empirical' theoretical principles or concepts. It is difficult to see how these conceptual activities would fit into the traditional representational picture.

This very indirectness of models as regards real world systems has been pinpointed by Weisberg (2007) and Godfrey-Smith (2006), who consider it as *the* distinctive characteristic of modelling. According to Weisberg and Godfrey-Smith, modelling can be viewed as a specific theoretical practice of its own that can be characterized through the procedures of indirect representation and analysis that modellers use to study the real-world phenomena. With indirect representation they refer to the way modellers, instead of striving to represent some real target systems directly, rather construct simple, imagined model systems to which only a few properties are attributed. They argue that in modelling, models come first in the sense that they are constructed and analysed before the relationship between the model and any target system is assessed, "if such an assessment is necessary" (Weisberg 2007, 209). Thus modellers may go on studying model systems without too much explicit attention to their relationship to the world, which makes models independent from any real target systems.

But how, then, are models as independent objects able to give us knowledge? Whereas Godfrey-Smith evokes the "*effortless* informal facility" with which we can assess similarities between imagined and real world systems, Weisberg refers to the notion of representation. But reverting to representation, not to mention mere similarity, would take us back to the problems discussed above. In contrast, what we find the most important point in viewing models as independent things is that it enables us to appreciate their functional characteristics, that is, the different purposes for which they are used in scientific practice. This gives us, we suggest, a clue to how to appreciate the epistemic properties of models from another perspective than that provided by representation.

In stressing the importance of the functional characteristics of models for their epistemic value, we follow Morrison and Morgan (1999) who point out that we learn from models by constructing and manipulating them. Also Morrison and Morgan consider models as independent entities. On their construal, models gain their (partial) independence from theory or data, since besides being composed of both theory and data, models typically also involve "additional 'outside' elements" (1999, 11). Boumans (1999), in turn, goes on to disentangle models from the theory-data framework altogether. In his view models are independent things constructed from as heterogeneous "ingredients" as analogies, metaphors, theoretical notions, mathematical concepts, mathematical techniques, stylised facts, empirical data and even relevant policy views. In consequence, part of the justification of a model is built into it through the initial justification of the ingredients that are "baked" into a model.

Taken together, the aforementioned approaches suggest that models should be addressed as independent things with some initial built-in justification and from which we learn by manipulating them. Building on this approach, we wish to go even one step further: We suggest conceiving of models as *concrete objects*, constructed for certain *epistemic aims* making use of various representational means and whose cognitive value derives largely from our *interaction* with them (Knuuttila 2008, Knuuttila and Merz 2009). Consequently, scientific models can be considered as multifunctional *epistemic tools* (Knuuttila 2005, Knuuttila and Voutilainen 2003). The point of stressing the concrete artefactual nature of models follows from our focus on the epistemic functioning of models. In other words, if our aim is to understand how models enable us to learn from the processes of constructing and manipulating them, it is not sufficient that they are considered as independent; they also need to be concrete in the sense that they must have a tangible dimension that can be worked on. This concreteness is provided by the material embodiment of a model: the concrete representational media through which a model is achieved gives it the spatial and temporal cohesion that enables its manipulability.

Our suggestion of approaching models as *epistemic tools* involves several largely novel features:

Firstly, it conceives of models as concrete constructed objects, which are expressed by external representational means. The construction of these epistemic artefacts both "affords and limits" scientific reasoning. While the idea of models as concrete objects dates back to the nineteenth century, what is new about our conception is that it makes no in principle distinction between material and abstract models. When working with abstract objects we typically also assemble and manipulate external representational means such as diagrams or equations. Certainly, building a material model or writing down a set of equations to describe a hypothetical system amount to very different kinds of epistemic activities, but from our perspective their differences derive partly from the representational means used, which are important for how we gain knowledge through models. More generally, the wide variety of representational means modellers make use of i.e. diagrams, pictures, scale models, symbols, natural language, mathematical notations, 3D images on screen) all *afford* and *limit* scientific reasoning in their characteristic ways. For instance, pictures or graphs both "afford" different kinds of reasoning than linguistic expressions or mathematical equations (see e.g. Knuuttila 2005, Vorms forthcoming, Kress & van Leeuwen 2001). Consequently, in putting forth the notion of models as epistemic tools we do not deny that models were representations in the sense of being entities that are constructed by making use of representational tools. Rather, as our account does not focus on the general features of a supposed representational relationship between a model and a target system it makes room for considering the *actual representational means* with which scientists go on representing.

The other part of the affording and limiting nature of models can be attributed to their *constrained design*. This is one of the main roles of the idealizations, simplifications and approximations made in modelling. Through these epistemic strategies modellers seek to design their models in such a way that it renders their initial problem more accessible and workable and helps them to tackle it in a more systematic manner. Thus models provide by their construction external aids for our thinking, something that cognitive scientists have approached in terms of the notion of *scaffolding*. According to them external representational scaffolding both narrow the space of information search by localizing the relevant features of the object in a perceptually salient and manipulable form, and enable further inferences by making the previously obscure or scattered information available in a systematic fashion (see e.g. Larkin and Simon 1989, Clark 1997, Zhang 1997, Sterelny 2004). Consequently, idealizations, simplifications and approximations need not be considered distortions or shortcomings of models since what might appear to be misrepresentation could also be part of a purposeful representational strategy.

Secondly, our approach puts to the fore the *evolving* nature of model building: Models are typically constructed step-by-step which follows from considering models as concrete artefacts, the epistemic value of which follows from our interaction with them. We do not, however, contend that models boil down to their concrete embodiment in some representational modes and their associated media, neither that their epistemic value follows from this dimension alone. Rather, we suggest that the concrete embodiment of a model (whether symbolically, iconically or three-dimensionally rendered), draws together and integrates, in each stage of its development, the various empirical, theoretical and conceptual dimensions of its construction. In fact, as we will show in the case of the Carnot model, modelling typically involves a theoretical (re)description of the target *phenomenon* as well as the development of *theoretical principles* and *scientific concepts*. The model in the process of its construction functions as an integrating tool as well as a scaffold for further scientific reasoning. In this way the model functions also as a tool of its own development.

Thirdly, as we do not wish to start out by assuming that models necessarily represent some external target systems, what are, then, their targets? From our perspective, models are purposefully crafted devices the aim of which is to provide answers to some pertinent scientific problems. Thus the starting point of modelling is more often than not the scientific question to be dealt with, rather than the attempt to represent a certain real target system more or less accurately.[iv] That these two tasks do not necessarily boil into one is due to the fact that scientific questions are usually general in nature, owing to the previous body of knowledge and pending scientific and practical problems. The proponents of the representational approach might try to take this into account by arguing that models strive to represent only some selected aspects of real target systems. Yet, from our perspective, this kind of reply amounts to putting the cart before the horse. Modelling is a way to inquire what is relevant for a certain problem, taking into account that a well-formulated question is already an epistemic achievement. We might arrive at a certain representational relationship as a result of our inquiries, but to assume that we have gained knowledge first when such a relationship is established is to assume too stringent criteria as regards the actual scientific knowledge. Indeed, approaching knowledge in the representationalist fashion may itself be questioned.

Consequently, and fifthly, our notion of models as epistemic tools pays heed to the different, though connected, connotations of the word “tool”. From our perspective models are artefacts, and like other human-made things they are made in view of certain purposes and thus function as tools. Approaching models through the tool-metaphor pays attention to the diverse tasks of models in science, such as prediction, design of experiments, theory development and scientific understanding. These tasks may be practically motivated: In engineering sciences, for instance, models are typically developed for the purposes of producing, controlling, or preventing some properties of materials or behaviour of processes and devices. On the other hand we wish to align our account with those developments in the philosophy of mind and language as well as in cognitive sciences that try to seek alternatives for *representationalism*. Often those accounts refer to the notion of a tool in order to avoid understanding mental content or linguistic meaning in representational terms. The basic idea is that instead of approaching knowledge as a repository of representations that reproduce accurately, i.e. stand truthfully for, mind-independent reality, one might approach language, theories or concepts as things that humans use in order to do certain things or to achieve certain results or effects. This does not mean that models could not represent some real systems – quite to the contrary, they are often *used to represent* some real systems, as the pragmatists stress. However, what seems problematical and even presumptuous as regards the actual scientific practice is the attempt to ground knowledge claims on some privileged representational relationship between a model and some real target system that prevails independently of our cognitive activities.

One crucial task relegated to scientific representation is that of providing a connection between our theoretical representations and the world. However, the very problem of establishing the link between our representations and the world is created by some representationalist presuppositions. The crucial difficulty with the representationalist theory of mind is that internal representations are supposed to stand for something else, but there is no access to this something else except via another representation. The problem of scientific representation is but one variant of this basic problem. However, from our perspective scientific models are not freely floating objects in need of being linked to the real world: they are already connected to the real world by way of the scientific questions and the existing body of theoretical and empirical knowledge, which both motivate and enable their construction. For instance, in the engineering sciences scientists characteristically build models for the purposes of imagining and reasoning about how to improve the performance of the devices, processes or materials of interest. Thus what is known already about these devices, processes and materials forms the background of modelling and the resulting models are intermixtures of fictional and factual: They involve imaginable properties and processes, and they incorporate measurable physical variables and parameters that allow the coordination of the model to the empirical data.

In the following section we will exemplify the aforementioned aspects of our conception of models as epistemic tools by studying the Carnot model of the heat-engine. We highlight the process of modelling paying attention to the aim of the model, the way the original problem motivating the model construction was translated into a phenomenon to be accounted for, and how the model in the course of its development became a tool of its own development.

4. The Case of the Carnot Model of a Heat-Engine[v]

How else can models give us knowledge than via a representational relation between a model and some real target system? In what follows we will analyse the development of the Carnot model in order to illustrate how conceiving of models as epistemic tools helps us to approach their epistemic value from a novel perspective. The key to the epistemic productivity of models lies in the activity of modelling, through which scientists develop a model step-by-step, building into it new aspects by which the content of the model becomes richer and more advanced. This may also lead to the construction of other models (see Peschard, this issue). In each modelling step the model functions as an epistemic tool that ‘affords and limits’ its own further development, functioning also as a springboard for the formulation of theoretical principles and concepts.

Our reconstruction of Carnot’s construction of the model of a heat engine is primarily based on the line of his argument presented in *Reflexions on the Motive Power of Fire and on Engines fitted to develop that Power* (Carnot 1986, [1824]). However, the modelling steps distinguished below do not necessarily present a sequential order – often these different aspects are modelled in a mutual interaction. As to the generalizability of our case, we do not claim that all the following steps always are present in modelling, nor, that any other steps could not be distinguished in this or other modelling examples.

Regarding the importance of the actual representational means used in model construction, it is worth mentioning that Carnot cast his model in terms of discursive reasoning by making assumptions and formulating principles in ordinary language. He did not have at his disposal such representational means as the well-known P-V diagram that was invented by Benoît Paul Émile Clapeyron only ten years after Carnot published his *Reflexions*, which enabled their successors to draw new conclusions from the Carnot model.[xv] Nevertheless, we claim that the conceptions Carnot developed by means of linguistic formulations enabled him to build the model step-by-step, in which process the concrete discursively expressed model indispensably functioned as an epistemic tool in its own making.

Step I. Articulating the epistemic aim

The engineering sciences usually start from questions related to practical problems and applications such as, for instance, the problem of how the functioning of a device can be improved. The problem that the French physicist and engineer Sadi Carnot got interested in was how to improve the performance (or efficiency) of heat engines. In his *Reflexions*, he writes that, "The study of these engines is of the utmost interest [because] their importance is immense, and their use is increasing daily." (ibid p. 61). Carnot was not interested in the 'trial-and-error' approaches familiar to engineering of his days. Instead, he aimed at a theoretical answer to a fundamental question about the performance of heat engines:

> "The question whether the motive power of heat [i.e. the useful effect that an engine is capable of producing] is limited or whether it is boundless has been frequently discussed. Can we set a limit to the improvement of the heat-engine, a limit which, by the very nature of the things, cannot in any way be surpassed? Or conversely, is it possible for the process of improvement to go on indefinitely?"(ibid p. 63)

This fundamental question illustrates a kind of general *epistemic purpose* common to the engineering sciences, to wit, finding out the fundamental limit to the improvement of the desired (or to the minimizing of the undesired) capacity of a technological artefact via theoretical understanding of the very nature of this capacity. For this purpose a model of an ideally functioning technological artefact is constructed, which affords and limits scientific reasoning about possible, or hypothetical, interventions with the real technological devices (e.g. the real heat engines). Consequently, the Carnot model can be conceived of as an epistemic tool for reasoning about why, on the one hand, certain losses cannot be avoided, and how on the other hand, one could aim to minimize these losses of the heat engine. From this perspective the aim of modelling is not primarily that of *representing* some real target-system more or less accurately, but rather producing a hypothetical device that meets some specific epistemic aim. For instance, drawing and describing the mechanical workings of an existing heat-engine (in engineering) can be considered as a representation of some real target system [vi], but as we will show, what Carnot established was something quite different. The characteristic step-by-step construction of a scientific *model* is not due to the attempt to represent successively more in detail the different aspects of some real target system, but it rather reflects how the theoretical principles and theoretical conceptions develop as the model gets more sophisticated and how the model in each consecutive phase enables its further construction.

Step II. Discerning the target phenomenon

The purpose of Carnot's model was to give theoretical understanding about natural or fundamental limits of the performance of heat-engines. In order to get an intellectual grip on the problem, the problem had to be conceptualized in such a way as to make it cognitively accessible. This often involves conceptualizing some phenomenon in a new, not necessarily obvious manner.

How is the phenomenon then (re)conceptualized? Generally, developing a scientific model for a technological artefact such as the heat-engine, involves conceiving of its functioning in terms of particular *physical phenomena* that produce its proper or improper functioning. (This by the way shows also that the engineering sciences proceed in theorizing in the same way as the physical sciences). Hence Carnot assumed that, "in order to grasp in a completely general way the principle governing the production of motion by heat, it is necessary to consider the problem independently of any mechanism or any particular working substance." (ibid. p. 64) Hence, Carnot conceived of the functioning of the heat-engine, not primarily in terms of its mechanical working[vi], but as a device that produces *motion by heat*, which is the target phenomenon to be modelled.

Step III. Scientific conception of the target phenomenon

We suggest that developing a conception of the target phenomenon in order to make it scientifically accessible, is an integral part of modelling. The target phenomenon is not simply observed in a straightforward manner, but it is typically co-constructed along with the model. This idea about the construction of a target phenomenon deviates from representational understanding of the target phenomenon, which assumes it to stand apart from the model. Contrary to the idea that physical phenomena lay waiting for the perceptive scientist who inventively picks it up for further examination through modelling or other means, we emphasize that the target phenomenon is not "ready-made". In this we partly follow Massimi (2008) who has argued that physical phenomena are not "empirical manifestations of what is there". As an alternative, she defends the view that, "phenomena are conceptually determined appearances". Although we will not adopt Massimi's vocabulary of appearances versus phenomena, we agree with her that a description of a scientifically accessible phenomenon is not a straightforward, 'conceptually bare' representation of what there is. Instead, it has a built-in conceptual content. In other words, when scientists discern a phenomenon, they do this by bringing an occurrence ('appearance') either under already existing scientific concepts, or, sometimes, under a newly invented scientific concept. An occurrence to which no scientifically significant conceptual content has been attributed usually remains cognitively inaccessible, and is not even recognized as a phenomenon in need of scientific explanation. Thus while Carnot most probably started from empirical knowledge concerning the functioning of heat engines, the starting point of his modelling exercise is the formulation of the target phenomenon in terms of heat *causing* the production of motive power. It should be also noted that in proceeding this way Carnot was already building on the older concepts, too, such as the caloric concept of heat.

Carnot's new scientific conception of how heat produces motive power, can be summarized as follows: Motive power is produced by *transfer* (i.e., by *transportation* rather than by *transformation*) of heat from a hot to a cold body. This conception of the phenomenon is built on a conception of the nature of heat, which in Carnot's days was conceived of as a material fluid called caloric.

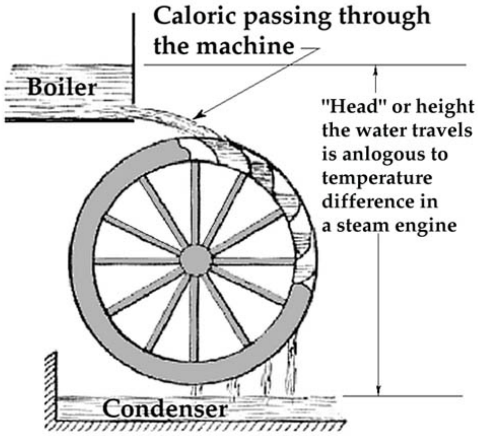
> "the production of motive power in a steam engine is due not to an actual consumption of caloric, *but to its passage from a hot body to a cold on*. It is due, in other words, to a restoration of the equilibrium of caloric after that equilibrium has somehow be disturbed ..." (ibid p. 65, his italics)

Carnot does not spell out his notion of caloric in *Reflexions.*[vii] [viii] A generally accepted account of caloric in those days was presented, for instance, by Dalton (1842, 1808), who stated:

> "The most probable opinion concerning the nature of caloric, is, that of its being an elastic fluid of great subtility, the particles of which repel one another, but are attracted by all other bodies. ..." (Dalton, 1842, p.1; first edition, 1808).[ix]

Carnot's conception of how heat produces motive power involved an analogy to how water-wheels produce motive power, as proposed by his father, Lazare Carnot, who had worked on a theoretical understanding of their efficiency (see Figure 1). According to this analogy, heat (caloric) spontaneously flows from high to low temperature, similar to how water flows from high to low levels, producing motive power (the turning of the water-wheel, *c.q.*, the movement of the piston of a heat engine). Similar to water-flow, caloric fluid flows from a hot to a cold body without being transformed or consumed itself. This conception implies that no heat is consumed in a cycle of the heat engine (in which gas is heated, then expanded – producing motive power, cooled, compressed, and heated again); *i.e.*, the *quantity* of heat in this cycle remains the same.[x]

In this way, Carnot developed a scientific conception of the target phenomenon by bringing it under a conception of how caloric produces motive power thus making it scientifically accessible.



Carnot's steam engine/waterwheel analogy

Figure 1. Sadi Carnot's conception of how heat flowing through a heat engine produces motive power, which is analogous to how water flowing through a water-wheel produces motive power.[xi]

Conceiving of the target phenomenon in terms of transporting caloric which is supposed to be carried around by steam cycling through the engine, enables Carnot to explain the functioning of the steam engine:

> "So what exactly happens in a steam engine of the kind now in use? Caloric produced in the furnace by combustion passes through the walls of the boiler and creates steam, becoming in a sense part of it. The steam bears the caloric along with it, transporting it first into the cylinder, where it fulfills a certain function, and then into the condenser. There, the steam is liquefied by contact with the cold water it encounters. In this way, at the end of the whole process, the cold water in the condenser absorbs the caloric produced by the initial combustion: it is heated by the steam just as if it had been in direct contact with the furnace. The steam serves simply as a means of transporting the caloric, ... we are considering the movement of the steam is put to use." (Carnot, 1986, [1824], p. 64)

This scientific conception of the target phenomenon already affords and limits the further modelling of the functioning of the heat engine and studying its theoretical limits.

Step IV. Conception of relevant (theoretical) principles

Carnot proceeded in his modelling endeavour by introducing propositions and principles that relate the conception of the transport of heat (caloric) and the production of motive power, to relevant measurable parameters such as temperature, volume, and compression or expansion of the gas in the steam engine. In this way he also tied the scientific aspects that were built into the model to data that can be observed or measured in the real systems.[xii] Some of these principles are definitions (e.g., a), others are experiential or experimental (e.g., b, c), and yet others theoretical (e.g. d, e, f, g, h). His development of these propositions and principles is reconstructed and summarized in the list below (ibid pp. 64-67):

> Firstly, Carnot presents a definition of heat engines that draws on his conceptualization of it:
>
> a. The heat-engine is any engine that is driven by caloric.
>
> He articulates also an experiential principle, which is important for the consecutive modelling exercise:
>
> b. Equilibrium restores wherever a difference in temperature exists.
>
> Additionally, there is an experimentally well-known fact that:
>
> c. The temperature of gaseous substances rises when they are compressed, and falls when they are expanded.[xiii]

This experiential and experimental principles are developed further by bringing them under the scientific conception of heat (caloric), resulting in several theoretical principles (d, e, f, g):

d. Caloric will always flow from a hot body to a cold body until the two bodies have the same temperature, by which equilibrium is restored.

Therefore:

e. In steam engines motive power is produced by the re-establishment of the equilibrium of caloric, not by consumption of caloric,

and:

f. whenever there is a difference in temperature, motive power can be produced,

while the converse is also true:

g. wherever there is power which can be expended, it is possible to bring about a difference in temperature and to disturb the equilibrium of caloric. [xiii]

From these principles Carnot infers 'an obvious' principle that

h. Heat can only be a source of motion in so far as it causes substances to undergo changes in volume or shape.

What this analysis shows is that Carnot built into the model several elements by stepwise conceptualization: he drew on some experiential or experimental knowledge, and enriched this knowledge by bringing it under relevant concepts creating thus a new way of imagining what is going on in heat engines.

Step V. Conception of a hypothetical device:

A simplified picture of the real heat engine consists of a furnace, a boiler, a steam containing cylinder closed with a movable piston, a condenser, and a reservoir of cold water.[vi] The explicitly articulated principles of the model together with this simplified picture of the steam engine enabled Carnot to imagine a hypothetical device that produces motive power by heat. This model functioned as an epistemic tool that guided and constrained its further development.

To start with, Carnot asked the reader to "imagine" two bodies, A and B (the temperature of A is higher than B), to which heat can be added or from which it can be taken away without effecting any change in their temperature. A and B will act as two infinite reservoirs of caloric. Next, he envisaged a hypothetical device consisting of three operations (ibid pp. 67-68):

"If we wish to produce motive power by conveying a certain amount of heat from the body A to the body B, we may do this in the following way":

(i) Take some caloric from the body A and use it to form steam. In other words, use the body as if it were the furnace. It is assumed that the steam is produced at precisely the temperature of the body A.

(ii) Pass the steam into a vessel of variable volume, such as a cylinder fitted with a piston, and then increase the volume. When the steam is expanded this way, its temperature will inevitably fall. Suppose that expansion is continued to the point where the temperature becomes exactly that of body B. [use of principle g].

(iii) Condense the steam by bringing it into contact with B, and, at the same time, subjecting it to a constant pressure, until it is totally liquefied. In this way, B fulfils the role of the injection water in a normal engine. (See ibid pp. 67-68).

Step VI. Introduction and use of an abstract concept: reversed processes

However, by the operations described in i, ii, iii, only half of the cycle through which the hypothetical device must go, has been constructed. For closing the cycle, the 'liquefied steam' at temperature B (in iii) must be brought back to temperature A. This could easily be achieved by heating the 'liquefied steam' to temperature A, as it may happen in a real heat-engine. However, at this point, Carnot made a brilliant conceptual leap by introducing the idea of a *reversed* process. Carnot stated that there is no reason "why we should not form steam with caloric from the body B and at the temperature of B, compress it so as to bring it to the temperature of A, continuing the process of compression until complete liquefaction takes place" (ibid p. 68). Carnot thus conjured up a sequence of processes that bring about a closed cycle, which is achieved by reversing the process so that the steam in the cylinder is brought back to its initial state.

We can easily follow Carnot in reasoning on an abstract level that operations i, ii, iii, could possibly be reversed. However, what is remarkable about his conception is that this reversal does not draw on concrete experiential or experimental knowledge. In contrast, experiential and experimental knowledge of that time would have hindered any such reasoning, for it was not easy to imagine, for instance, that steam could be formed from water at temperature B by a cold body at temperature B (which is what actually happens in refrigerators). The introduction and use of this concept is a good example of how creative reasoning works. Such reasoning is made possible by bringing knowledge that has been systematically and explicitly brought together in the model under an abstract concept (reversed processes), leading in turn to new conceptions (forming steam at a low temperature by the transfer of caloric to it).

Nevertheless, the *hypothetical* idea of the reversible process raises the question of why we should believe that it is *physically* possible as well. Indeed, in order to explain how this would work physically, the conception of the reverse process needs to be fleshed-out much further, which is what happened in Carnot's further modelling.

The imagined possible operations (i, ii, iii and reverse) guided Carnot in constructing a closed cycle that produces the maximum amount of motive power. First, Carnot used the notion of the reversed process to imagine a cycle in which in the first sequence of operations (i, ii, iii), motive power is produced and at the same time caloric is transferred from body A to body B, while in the reverse operations, exactly the same amount of motive power is expended and caloric returns from B to A. In this cycle no net motive power is produced, neither is there any net transfer of caloric from A to B. Carnot then argued that if it were possible to make caloric to yield a greater amount of motive power in the first sequence (i, ii, iii), we should have a case of motive power being created in unlimited quantities without the consumption of caloric, which contradicts the idea that perpetual motion is impossible for mechanical processes. By this reasoning, Carnot showed that in order to produce net motive power in a closed cycle, there must be a transfer of heat from A to B in a cycle. (ibid. 69)

Secondly, this idea of a closed cycle guided in specifying the performance (efficiency) of the hypothetical device as the ratio between the quantity of motive power developed in a complete cycle of operations and the amount of heat transferred from A to B.[xiv] The motive power developed is defined as the product of the volume and the difference of its pressure at the expansion of the gas (sequence i, ii, iii) minus this product at the compression of the gas. The total amount of heat needed is the amount of heat transferred from A to B (ibid 98)

Step VII. Deriving principles for producing the maximum amount of motive power, i.e., principles for avoiding losses

Carnot's next question was how the *maximum* amount of power could be obtained, and what 'maximum' in this context means. The model of the hypothetical device in its current state suggested that it meant minimizing the causes of the loss of heat. Further development of the model thus required accounting for the possible causes of loss of the hypothetical device. Again, the model in its current state, in particular the formerly stated principles (a-h), enabled Carnot to develop some additional principles that explained losses (and the avoidance of losses). These principles are the following (ibid p. 70):

i. Since any process in which the equilibrium of caloric is restored can be made to yield motive power, a process in which the equilibrium is restored without producing power must be regarded as representing a real loss.

Reflecting on this latter point, Carnot concluded:

j. Any change in temperature that is not due to a change in the volume of a body is necessarily one in which the equilibrium of caloric is restored profitlessly.

Hence:

k. The necessary condition for the achievement of maximum effect is that the bodies used to produce motive power should undergo no change in temperature that is not due to a change in volume**.**

However:

l. When a gaseous fluid is rapidly compressed, its temperature rises; and when, on the other hand, it is rapidly expanded, there is a fall in temperature.

What becomes obvious as one looks at these latter principles (i-l) is that Carnot had to construct a cycle in which a change in a temperature without a change in volume (j) was avoided, as well as rapid compression or expansion (l).

Step VIII. Constructing the ideal heat-engine

At this point, Carnot' s model encompassed a hypothetical device (that consisted of a cylindrical vessel closed with a movable piston that encloses a constant amount of steam – in which the gas can be either thermally isolated, which means that there is no transfer of caloric, or contacted with body A at a constant high temperature that acts as a source of caloric [heat], or with a body B at a constant low temperature that acts as sink of caloric [heat] – which goes through cycle i, ii, iii, and reverse), as well as knowledge by which he could construct the possible operations (i.e., principles a-h), and knowledge about the causes of loss (i.e., principles i-l). Carnot used this model as an epistemic tool for constructing a cycle that produces the maximum amount of motive power. That is, he used the model at this stage for constructing a hypothetical device called the ideal heat engine.

As we already have illustrated in the former section (Step VI), Carnot's approach entailed using the model for imagining different kinds of possible operations with the hypothetical device. At this point, he aimed at constructing a hypothetical cycle that avoids losses, which, as he knew by now, had to be constructed in such a way that the problem of "restoring caloric profitlessly" was avoided by preventing that "changes in temperature occur that are not due to a change in volume". In the further construction of the hypothetical device he decided to use gas instead of steam, which simplified the hypothetical device to the extent that the processes of condensation and evaporation in a cycle could be neglected.

Carnot imagined, for instance, how the temperature could be changed by withdrawing or supplying caloric without changing its volume. According to the principle j, this is an operation that causes loss, which therefore must be avoided in the ideal heat engine. Carnot argued that it would be equally possible to withdraw caloric during the process of compression in such a way that the temperature of the gas would remain constant, which implies that the rise of temperature that would be due to rapid compression (principle l) can be avoided. Likewise, if the gas is expanded, its temperature can be prevented from falling if we supplied to it an appropriate quantity of caloric. (ibid. 72). We claim that imagining such operations results from using the model as an epistemic tool, and they could not have been derived from mere experience with real steam engines.

By this way of reasoning, Carnot proposed a hypothetical cycle of four operations that is supposed to produce the maximum amount of motive force, arriving thus at the ideal heat engine. The description of the four operations of the ideal heat engine refers to the schema of the cylinder with piston that can be contacted with body A and B (Figure 2):

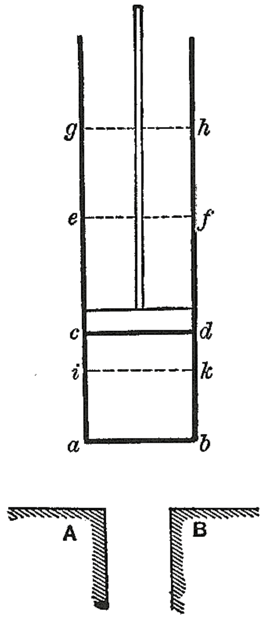


Figure 2. (ibid. p. 114) Axial cross section of the hypothetical device, which is part of the Carnot model of the heat-engine. In this diagram, *abcd* is a cylindrical vessel, *cd* is a movable piston, and *A* and *B* are constant–temperature bodies that act as source or sink of caloric. The vessel may be placed in contact with either body or removed from both (as it is here).

(1) The air is placed in contact with the body B. It is then compressed by returning the piston from its position *gh* to *cd*. During this process, the air maintains a constant temperature, since it remains in contact with B and gives up caloric to it.

(2) The body B is taken away, and the compression of the air is continued. Since the air is now isolated, its temperature rises. Compression continues until the temperature of the air reaches that of the body A, by which time the piston has moved from the position *cd* to *ik*.

(3) The air is placed once again in contact with the body A, and the piston returns from *ik* to *ef*; the temperature remains constant.

(4) A is removed, so that the air is no longer in contact with any body that can act as a source of caloric. But the piston continues to move, rising from the position *ef* to *gh* The air expands without absorbing caloric, and its temperature falls. Let us *suppose* that the temperature continues to fall until it is equal to that of B, whereupon the piston stops at the position *gh*. (see ibid pp. 74-75).

This description of the cycle drew, as we have seen, on the previously developed conceptions of possible operations.[xv] Hence, Carnot's model of a heat engine finally covered the epistemic aim expressed in terms of the target phenomenon, different pieces of knowledge as presented in principles a-l, and the final formulation of the ideal heat-engine, which is a hypothetical device that is supposed to produce motive force at minimum loss (i.e., the maximum amount of motive force that can be produced by heat-engines).

Step IX. Relating the model to the empirical data and producing new knowledge through it

Carnot included in his model variables, such as the pressure and the volume of the gas, that hook the model to the observable or measurable world. This may suggest that in the remainder of his *Reflexions* (ibid pp. 78-112), Carnot aimed at testing the model by comparison of its predictions to empirical data. However, even if it were possible to build a device that resembles it (a gas-containing cylinder closed with a frictionless piston, alternately contacted to or isolated from the sources of heat), methods to quantify variables such as the temperature and the amount of heat (caloric) of gas had not been established.[xvi]

In other words, in *Reflexions*, testing and justification of the model does not take place by the comparison of the outcomes of the model to real data, but only by its construction (i.e., along the lines of steps I-VIII). This is not to claim that the results of models (i.e. the conclusions produced by using models as epistemic tools) are never put to empirical tests, neither that their justification merely consists of how they are constructed. Moreover, as we have seen, empirical findings are crucial for how models are constructed, and thus they play a role in their justification. Undeniably, new empirical findings, but also new theoretical insights, may lead to revisions of (some aspects of) the model – indeed, this is what happened to Carnot's model when the caloric theory of heat was substituted.

One important question is whether Carnot was actually able to use the ideal heat engine as an epistemic tool for reasoning about the real heat engines thus producing new knowledge concerning them. In the remainder of his *Reflexions* (ibid pp. 78-112) Carnot indeed used the model in this way: for instance, he discussed variables that could affect the efficiency of the device, such as the working substance (e.g., heat or steam or alcohol), the temperature of the bodies A and B, and their temperature difference. In this discussion he related the model to actually available measured data and phenomenological laws, thereby producing new physical knowledge of gasses, presented in propositions such as:

> "The motive power of heat is independent of the working substance that is used to develop it. The quantity is determined exclusively by the temperature of the bodies between which, at the end of the process, the passage of caloric has been taken." (ibid 76-77).[xvii]

and

> "The difference between the specific heats at constant pressure and the specific heat at constant volume is the same for all gases." (ibid 80).

Besides of finding out which factors affect the efficiency of the device, he also aimed to find ways for determining specific properties of gasses, in particular specific heats (the amount of heat needed to raise its temperature with one degree at either constant volume or at constant pressure), that would allow for making calculations on, e.g., the efficiency.[xviii] Finally, Carnot used the physical knowledge of gasses thus developed, together with available quantitative data, for estimating the heat transfer in a cycle per unit of motive force, and hence the maximum efficiency of the ideal heat engine. In these examples, the model enabled the production of knowledge that was not already contained in it by the use of it as an epistemic tool for reasoning about empirical data (of gasses) towards new conclusions concerning the properties of gasses and heat engines, respectively.

Summing up

In our discussion of the Carnot model we have aimed to show that the Carnot-model of the ideal heat-engine aims not to be a *representation* of actual heat-engines bearing some kind of obvious morphism or similarity with them – which would be the case, for instance, if it strove to depict the mechanical workings of them. Rather, the Carnot-model is better seen as a hypothetical engine affording reasoning in view of a certain purpose. In our reconstruction, we have focussed, firstly, on the crucial role conceptualization plays in modelling, and secondly, on how modelling consists of a co-construction of different mutually developing elements that are drawn together by the model. Such constructed elements include importantly (1) the epistemic aim of the model, usually related to a pending scientific question or problem; (2) the idealizations, abstractions and simplifications that make the problem manageable and workable; (3) the target phenomenon into which the original problem is translated; (4) the particular representational means with which the imaginary (or hypothetical) target system is represented; (5) the experiential and theoretical knowledge made use of in model construction; (6) the concepts, principles and conceptualizations, some of which may emerge in the process of modelling; and (7) the relevant observable or measurable parameters, which enable the coordination of the model with real systems. We claim that this intricate content of scientific models, which usually is fully understood only by the scientists working in the field in question, makes them to function as epistemic tools.

5. Concluding remarks

How do models give us knowledge? We have claimed that this happens through the construction and use of concrete artefacts, models that are embodied with different representational means. It seems to us that the key to the epistemic value of models is the process of modelling rather than some determinable representational relationship between a ready-made model and its putative real target system. We have shown with the example of the Carnot model of the heat engine that in the process of modelling, the model construction coincides with the creation (and discernment) of phenomena, and the development of theoretical concepts and principles. While the idea of such co-construction of phenomena, theoretical concepts and principles is by no means novel, it has not really entered the discussion on modelling in the philosophy of science. This may seem odd given that models are typically delegated the roles of scientific discovery and theory construction in the philosophical discourse. It seems to us that this neglect is explained by the prevailing representational approach. The aim of capturing the objective nature of real target systems (either by similarity or by a morphism of some kind) seems to be built into the representational approach. Such perspective is at odds with the idea of co-construction of the theoretical model and its target system.

While the pragmatic approaches to representation do not seek such an objective representational relation that would succeed to link the model and the target irrespective of the intended uses of the model, they also have focussed so far in the use of ready-made representations. To be sure, Giere (2006) and van Fraassen (2008) have pointed out the importance of the creation of phenomena as a part of their pragmatic approaches to representation. This does not save, however, the pragmatic accounts of representation from its deflationism, which implies that it cannot be expected to do any significant philosophical work in explaining in virtue of what we gain knowledge through models. We do not regard this, however, as an inherent deficiency of the pragmatic account of representation, since to expect a philosophical analysis of representation to ground our knowledge claims seems to us a mission impossible. Moreover, our critique against the representational account of models is not geared against the possibility of representation per se. We have rather argued that the representational approach to models does not succeed in delivering what it is commonly taken to establish, i.e. explaining in virtue of what do models give us knowledge.[xix]

While models make part of the class of cultural artefacts called representations, from our point of view establishing a *representational relationship* between a model and a part of the world is already a remarkable epistemic achievement. We have argued that the activity of modelling forges the link between the representational vehicle and our knowledge of the real world. The key to the mystery of representation lies in the process of co-construction, in which process the model, the phenomena, as well as the theoretical concepts and principles are developed concurrently. We have also suggested that the model as a concrete evolving object functions as a tool in integrating and guiding this process. In consequence, the model does not function as an epistemic tool because of some pre-determined representational relationship, but functions rather as a tool in establishing it – apart from functioning as a tool for various kinds of scientific reasoning. However, once a representational relationship becomes established, we tend to become forgetful of this co-constructive process by which it came into being. This is the point at which the philosophical wrestling with representation tends to start.

Literature

Ambrosio, C. (2007), *"Iconicity and Network Thinking in Picasso's Guernica: A Study of Creativity across the Boundaries".* Unpublished PhD thesis, University College London.

Bailer-Jones, D. M. (1999),"Tracing the Development of Models in the Philosophy of Science", in Lorenzo Magnani, Nancy J. Nersessian and Paul Thagard (eds.), *Model-Based Reasoning in Scientific Discovery*. New York: Kluwer, 23-40.

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

Bailer-Jones, D. M. (2009), *Scientific Models in Philosophy of Science*. Pittsburgh: University of Pittsburgh Press.

Bartels, A. (2006), "Defending the Structural Concept of Representation", *Theoria* 55, 7-19.

Boon, M. (2006), "How Science is applied in Technology", *International Studies in the Philosophy of Science* 20(1): 27-47.

Boon M. (2009), "Understanding according to the Engineering Sciences: Interpretative Structures", in Henk De Regt et.al. (eds.), *Scientific Understanding: Philosophical Perspectives*. University of Pittsburgh Press, 249-270.

Boon, M. and T. Knuuttila (2009), “Models as Epistemic Tools in Engineering Sciences: A Pragmatic Approach”, in: *Handbook of the Philosophy of Technological Sciences*, Anthonie Meijers (ed.), Amsterdam, Elsevier Science. 687-719.

Boumans, M. (1999), “Built-In Justification”, in Mary S. Morgan and Margaret Morrison (eds.), *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press, 66-96.

Carnot, S. (1986, [1824]), *Reflexions on the Motive Power of Fire*. Translated and edited by Robert Fox, New York, Manchester University Press.

Clark, A. (1997), *Being There: Putting the Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Chang H. (2003), “Preservative Realism and Its Discontents: Revisiting Caloric”. *Philosophy of Science* 70: 902-912.

Chang H. (2004), *Inventing Temperature, Measurement and Scientific Progress*. Oxford University Press.

Clausius, R. (1865), *The Mechanical Theory of Heat – with its Applications to the Steam Engine and to Physical Properties of Bodies*. London: John van Voorst, 1 Paternoster Row. MDCCCLXVII

Clausius, R. (1899, [1850]), “On the Motive Power of heat, and on the Laws which can be Deduced from it for the Theory of Heat”. In: *The Second Law of Thermodynamic. Memoirs by Carnot, Clausius and Thomson*. New York and London, Harper & Brothers Publishers.

Da Costa, N. C. A. and S. French (1990), “The Model–Theoretic Approach in the Philosophy of Science”, *Philosophy of Science* 57: 248–265.

Dalton, J. (1842), *A new System of Chemical Philosophy*, Part 1, (second edition), London, John Weale, Architectural Library, High Holborn.

Erlichson, H. (1999), “Sadi Carnot, ‘Founder of the Second Law of Thermodynamics”, *European Journal of Physics* 20: 183-192.

Fox, R. (1986), *Sadi Carnot, Reflexions on the Motive Power of Fire – A Critical Edition with the Surviving Scientific Manuscripts*. Manchester University Press, Lilian Barber Press, Inc.

French, S. (2003), “A Model-Theoretic Account of Representation (Or, I Don’t Know Much About Art…but I Know It Involves Isomorphism)”, *Philosophy of Science* 70:1472-1483.

French, S. and J. Ladyman (1999), “Reinflating the Semantic Approach”, *International Studies in the Philosophy of Science* 13: 103-121.

Frigg, R. (2002), *Models and Representation: Why Structures Are Not Enough*. Measurement in Physics and Economics Discussion Paper Series. London: London School of Economics.

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

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

Giere, R. N. (1999), *Science without Laws*. Chicago and London: The University of Chicago Press.

Giere, R. N. (2002), “Scientific Cognition as Distributed Cognition”, in Peter Carruthers, S. Stich and M. Siegal (eds.), *Cognitive Bases of Science*. Cambridge: Cambridge University Press.

Giere, R. N. (2004), “How Models Are Used to Represent Reality”, *Philosophy of Science* (Symposia) 71: 742-752.

Giere, R.N. (2006), *Scientific Perspectivism*. Chicago University Press.

Giere, R.N. (2010), “An Agent-Based Conception of Models and Scientific Representation”, *Synthese* 172:269-281.

Godfrey-Smith, P. (2006), “The Strategy of Model-Based Science”, *Biology and Philosophy* 21: 725-740.

Hacking, I. (1983), *Representing and Intervening*. Cambridge: Cambridge University Press.

Humphreys, P. (2004), *Extending Ourselves. Computational Science, Empiricism and Scientific Method*. Oxford: Oxford University Press.

Knuuttila, T. (2005), “Models, Representation, and Mediation”, *Philosophy of Science* 72: 1260-1271.

Knuuttila, T. and Merz, M. (2009), “An Objectual Approach to Scientific Understanding: The Case of Models”, in Henk de Regt, Sabina Leonelli and Karl Eigner (eds.) *Scientific Understanding: Philosophical Perspectives*. Pittsburgh: University of Pittsburgh Press, 146-168.

Knuuttila, T. and Voutilainen, A. (2003), “A Parser as an Epistemic Artefact: A Material View on Models”, *Philosophy of Science* 70: 1484–1495.

Kress, G. & T. van Leeuwen (2001), *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Arnold: London.

Kroes, P. (1995), “Technology as Science-Based Heuristics”, in J.C. Pitt (ed.): *New Directions in the Philosophy of Technology*. Kluwer Academic Publishers: 17-39.

Kuhn, T. S. (1958), “The Caloric Theory of Adiabatic Compression”, *Isis* 49(2): 132-140.

Larkin, J. and Simon, H. A. (1987), “Why a Diagram is (Sometimes) Worth 10,000 Words”, *Cognitive Science* 11:65-99.

Lynch, M. (1990), “The Externalized Retina: Selection and Mathematization in the Visual Documentation of Objects in Life Sciences”, in M. Lynch and S. Woolgar (eds.), Repre*sentation in Scientific Practice*. Cambridge, MA: MIT Press, 153-186.

Massimi, M. (2008), “Why There are No Ready-Made Phenomena: What Philosophers of Science Should Learn From Kant”, Royal Institute of Philosophy Supplement 83(63): 1-35.

Mattila, E. (2006), *Questions to Artificial Nature: A Philosophical Study of Interdisciplinary Models and their Functions in Scientific Practice*. Philosophical Studies from the University of Helsinki, 14. Dark: Helsinki.

Mendoza, E. (1961), “A Sketch for a History of early Thermodynamics”. Physics Today, 14 (2) 32-42.

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

Morrison, M. and M. S. Morgan (1999b), ”Models as Mediating Instruments”, in Mary S. Morgan and Margaret Morrison (eds.), *Models as Mediators. Perspectives on Natural and Social Science.* Cambridge: Cambridge University Press, 10-37.

Mäki, U. (2009). “MISSing the World: Models as Isolations and Credible Surrogate Systems, *Erkenntnis* 70: 29-43.

Nersessian, N. J. (2008), Creating Scientific Concepts, Cambridge, MA: The MIT Press.

Peschard, I. (the publication year of this issue), “Making Sense of Modeling: Beyond Representation”, *European Journal for Philosophy of Science.*

Portides, D. P. (2005), “Scientific Models and the Semantic View of Scientific Theories” *Philosophy of Science* 72: 1287-1298.

Sterelny, Kim (2004), “Externalism, Epistemic Artefacts and The Extended Mind”, in R. Schantz (Ed.). *The Externalist Challenge. New Studies on Cognition and Intentionality* (pp. 239-254) Berlin & New York: de Gruyter.

Suárez, M. (1999), “Theories, Models, and Representations”, in L. Magnani, N. J. Nersessian and P. Thagard (eds.), *Model-Based Reasoning in Scientific Discovery*. New York: Kluwer, 75-83.

Suárez, M. (2003), “Scientific Representation: Against Similarity and Isomorphism”, *International Studies in the Philosophy of Science* 17: 225-244.

Suárez, M. (2004), “An Inferential Conception of Scientific Representation”, *Philosophy of Science* (Symposia) 71: 767-779.

Suárez, M. (2010), Scientific Representation. *Blackwell’s Philosophy Compass* 5, 91-101.

Suppe, F. (1974), *The Structure of Scientific Theories*. Urbana, University of Illinois Press.

Suppe. F. (1989), *The Semantic Conception of Theories and Scientific Realism*. Urbana and Chicago, University of Illinois Press.

Van Fraassen, B. (1980), *The Scientific Image*. Oxford: Oxford University Press.

Van Fraassen, B. (2008), *Scientific Representation: Paradoxes of Perspective*. Oxford: Oxford University Press.

Vorms, M. (forthcoming), “Formats of Representation in Scientific Theorising”, to appear in Paul Humphreys and Cyrille Imbert (eds.), *Representations, Models, and Simulations*. Routledge.

Weisberg, M. (2007), “Who is a Modeler”, *British Journal for the Philosophy of Science* 58: 207-233.

Zhang, Jiajie (1997), “The Nature of External Representations in Problem Solving”, *Cognitive Science* 21: 179-217.

Acknowledgements

This research is supported by a grant from the Academy of Finland and the Dutch National Science Foundation (NWO).

Notes

[i] We take it that like scientific models also philosophical accounts are constructed for certain purposes. Accordingly, from our perspective, conceiving of models as epistemic tools aims to be an useful tool itself - for the philosophical purpose of finding out how models and modelling gives us knowledge.

[ii] Representation appears as an activity, too. But then representation is approached differently than from the representational perspective. It does not boil down to a relationship between a model and a target system, but rather concerns conveying some content with the help of some representational means and relevant conventions stipulating their uses.

[iii] Some other pragmatically inclined philosophers of science have tried to overcome the deflationist nature of the pragmatist account by adding to their analysis of representation a further stipulation concerning its success. Rather unsurprisingly, then, what was previously presented as an analysis of the representational relationship, i.e., isomorphism (van Fraassen 2008) or similarity (Giere 2010), is now suggested as a success criterion. As for isomorphism, it poses too stringent a condition on the success of representation in the light of

scientific practice. The case of similarity is trickier. On the one hand, it does not really supply any user-independent success criterion in that it is the users who identify the "relevant respects and sufficient degrees" of similarity. Giere (2010) admits this, arguing that the agent-based approach "legitimates using similarity as the basic relationship between models and the world".

[iv] In this respect our approach comes close to Mattila's work on simulation models as "artificial nature" constructed to answer some pertinent questions (Mattila 2006).

[v] In a previous article (Boon and Knuuttila, 2009) we have presented an extended, more technical analysis of Carnot's work. The approach in the present article is different as to (1) emphasis on the role of conceptualizing in modelling, in general, and (2) emphasis on the co-construction of the target phenomena, theoretical conceptions and model.

[vi] An animation that represents the mechanical working of a specific type of heat engine (the Newcomen steam engine) can be found under this link: http://en.wikipedia.org/wiki/Newcomen_steam_engine

[vii] Our basic reconstruction of Carnot's conception of heat has been taken from Clausius' (1864, 1899) *Memoirs on Carnot*.

[viii] Kuhn (1958) argues that Carnot followed Poisson who took from the caloric theory only the hypothesis that the heat content of a gas is a state function (which means that the heat content of an amount of gas is fully determined by the pressure and the temperature of the gas), and further, that at fixed pressure the caloric content is proportional to volume. Accordingly, Poisson developed a formula for the dependence of heat capacity on pressure. These assumptions and formula enabled Carnot to carry out calculations in the empirical part of his *Reflexions* (starting on ibid p. 78).

[ix] This concept of caloric is enriched with the idea that temperature is the density of caloric. In a further theoretical elaboration, it was postulated that caloric exists of two different states: sensible and latent. In its free state, caloric was conceived of as sensible, being able to affect the thermometer and our senses, whereas in its latent state, caloric is combined with matter and deprived of its characteristic repulsive force, thus being unable to effect the expansion of thermometric substances. This refinement of the caloric theory allowed for explaining e.g., that addition or withdrawal of (latent) heat causes a change of a state (e.g., melting, freezing, boiling, condensation, etc.) without change of temperature (cf. Chang, 2003 and 2004).

[x] Importantly, Carnot's presupposition of the conservation of heat, which was in accordance with the caloric theory of heat, was replaced by the idea of the conservation of energy (which is the first law of thermodynamics) by the work of his successors such as Thomson (also see Chang, 2004).

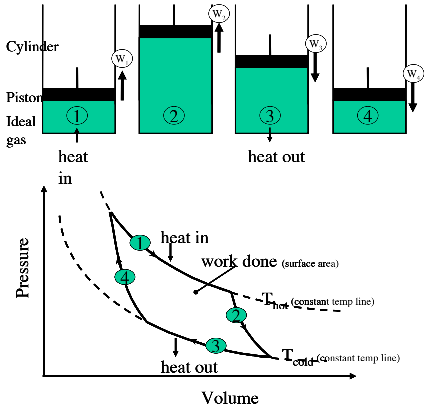
[xi] The Engines of Our Ingenuity is Copyright © 1988-2004 by John H. Lienhard.
http://www.uh.edu/engines/epi1958.htm

[xii] One should keep in mind that the measurements of several variables crucial to Carnot's model, such as temperature and (latent) heat, were not at all straightforward. Chang (2004) argues that the measurement and the conception of temperature had not been established in those days. To the contrary, Carnot's ideal heat engine played an important role in conceptualizing temperature such as how to conceive of a measure of one degree ("absolute" temperature) and how to conceive of measuring it.

[xiii] This experiential fact describes the so-called adiabatic heating and cooling. This phenomenon could very well have been taken as a counter-example to the theoretical idea that work *cannot* produce heat – an idea central to the caloric theory of heat. For, the sensible rise of temperature of a gas during compression suggests that heat is produced by expending power. Nevertheless, this phenomenon was (or has been brought) in accordance with the caloric theory by the distinction between sensible and latent caloric (also see former notes). Laplace, for instance, explained adiabatic heating and cooling by assuming that some quantity of latent heat is released when compressing the gas, producing an increased density of free (sensible) caloric that causes an increase of the temperature (see Kuhn, 1958, Mendoza 1961, and Chang 2003). This refined conception of caloric also may back-up theoretical principle *g*.

[xiv] Clearly, the transfer of heat cannot be directly measured. In his actual calculations, Carnot argues: “As for the heat which is used – that is, the heat transferred from A to B – this quantity is clearly that which is required to convert the water into steam...” (ibid 98). Note that the calculation of the efficiency in modern thermodynamics – which has abandoned the caloric theory of heat – uses the consumption of heat (i.e., the difference between the amount of heat entering and leaving the device) rather that the amount of heat that is transferred from body A to B.

[xv] This figure presents a modern conception of the Carnot engine. This conception has adopted conservation of energy rather than heat. It uses representational tools, such as the P-V diagram, that were only developed by Carnot’s successors.



[xvi] Chang (2004) shows that the definition and the measurement of temperature had not been settled; what is more, that establishing the measurement of temperature meats many, often entangled, practical and theoretical challenges causing that the development of the definition and measurement of temperature has been inextricably entangled with the development of thermodynamics. Significant to our case is how Thompson used Carnot’s conception of the ideal heat engine to define the interval of one degree of temperature (i.e., “absolute temperature”) as the amount that would result in the production of unit amount of mechanical work in a Carnot engine operating in that temperature interval (Chang, 2004, 182). Besides the fact that the ideal heat engine could not easily be operationalized, Thomson’s idea was abandoned when Carnot’s basic assumption concerning the conservation of the heat was rejected.

[xvii] Next, Carnot aimed to examine whether the fall of caloric from 100^0 to 50^0 yields more or less motive power than the fall of the same amount of caloric from 50^0 to 0^0., i.e., whether motive power is proportional to the difference in temperature.

[xviii] The difficulty of these properties of gasses was that the measurements had shown that they were not constants, but instead their values might depend on the temperature and the pressure of the gas. Moreover, it was not theoretically understood whether and how these values depended on them. Yet, Carnot’s reasoning in this part of the *Reflexions* sometimes is oblique and even remains incoherent. It not only involves empirical data and insights arrived at by the model, but also aspects of the caloric theory of heat which he usually does not spell out. Taking into account a reconstruction of the caloric theory that could back-up his reasoning (see former notes) rarely makes it more intelligible. Fox (1986, 26) argues that Carnot, while working on his *Reflexions* seemed to change his conception of heat from a caloric theory to the idea that “Heat is nothing but motive power”. In other words, Carnot’s conception of heat shifted from a caloric theory according to which heat is material (and in which heat is a conservative quantity that cannot be produced or consumed), to a dynamic theory according to which heat is a mode of motion (and were heat is no longer a conservative quantity but can be produced or consumed, e.g., by motive power). This may explain the sometimes inimitable manner in which Carnot reasons towards conclusions in the last part of his *Reflexions*.

[xix] One reading of the representational account of models, in line with the semantic version of it, is that the analysis of representation does not even try to explain how models give us knowledge but rather functions as a justificatory account. Thus, on that account, the analysis of representation is supposed to give us a criterion for

the success of the model. From our perspective the semantic account does not fare well even in this respect because the justification of a model gets largely built into it in the process of modeling. In this process the model is not just justified by the initial justification its ingredients might have, but also through the new theoretical and conceptual content born in the process of modeling.