mot1442 – Text 3

Theories and models

All views concerning the goals of science, or the function that science provides, share the idea that science is minimally a tool for making predictions. The instrumentalist conception of science sees prediction as the only goal of science. The term 'prediction' should be interpreted broadly, however: the predictions we use science for do not just concern future phenomena but equally phenomena or states-of-affairs that we could in principle verify directly to be the case right now but where the circumstances – lack of time, lack of resources, lack of adequate tools – prevent us from doing so. The prediction that 'This thing conducts electricity' on the basis of the claims that 'This thing is made of metal' and 'Everything made of metal conducts electricity' is of the latter kind. The conductivity of 'this thing' does not lie in the future, and we must have methods for establishing whether something conducts electricity, otherwise we could have no basis whatsoever for the claim that 'Everything made of metal conducts electricity'. Presumably, however, some condition – which could even be our laziness – prevents us from applying that method in the case of 'this thing'.

If we are instrumentalists regarding science, scientific knowledge need to consist of nothing more than laws in the sense of empirical generalizations, of the kind 'Everything made of metal conducts electricity'. The only element that makes these statements escape from direct empirical verifiability is their being universal statements, containing terms like 'everything', 'all', 'always', and so forth. This element requires us to apply some form of reasoning to arrive at their truth, but we have little difficulty in grasping what their truth consists in, giving that we have empirical access all other terms they contain, such as 'being made of metal', 'conducting electricity', 'having a pressure of so-and-so many kilograms per square centimetre', 'having a volume of so-and-so many litres', and so forth.

If, in contrast, the goals of science are seen as including explanation and understanding beyond furnishing predictions, then the content of scientific knowledge contains theories and models beyond mere empirical generalizations. Examples of such theories are classical mechanics, electrodynamics and quantum mechanics in physics, the biomolecular theory of DNA reduplication, DNA expression and protein synthesis in biology, the theory of rational producer and consumer behaviour in economics and more broadly every theory in psychology and social science that ascribes mental states or attitudes to people. The reasons why we need reasoning to arrive at the truth of such statements now become more variegated. Not only are many statements in these theories universal statements – according to classical mechanics *every* body undergoes an acceleration proportional to its mass when acted upon by a force – but they also contain terms to which we do not have direct empirical access – as for instance the term *force*. Apart from the difficulties we encounter in coming up with a form of reasoning that indeed gives us the truth of such statements, difficulties that were discussed in the chapter 'The validation of scientific claims', we now face an additional difficulty: what does the truth of such theoretical statements, and of theories in general, consist in?

A naive first answer to this question would be that the theory describes the world just as our empirical claims describe the world, but now including the parts that we can't observe. The universal markers 'every' and 'always' contained in any law represent patterns of causality or 'natural necessity' at work in the world and terms like 'force', 'energy', 'gene', 'desire' and 'expectations' contained in theories for various domains refer to non-observable things and describe their properties. This view, however, runs into a load of problems. Some of these are fairly technical and restricted to specific domains. To give just one example, the physical world as quantum mechanics describes it – that is, if we take that theory as describing part of the world – seems to be populated with things that mock our understanding of what it is to have a property, since these things sometimes have a property and also do not have that property, or are undecided concerning whether or not they have a property and behave a bit like having it and a bit like not having it. This puts serious doubt to the idea that this theory indeed describes the world at its deepest level. However fundamental these problems, they remain restricted to physics from the perspective of science in general. An objection against the naive answer that applies everywhere is that we know from the start that the theory does not truly describe the world. The world is of an incredibly larger complexity than any scientific theory portrays it. Take a particular theory from physics: the kinetic gas theory. It represents a gas as consisting of a great many extremely small particles that all have a mass and a velocity and all bump into each other and into the walls of the container (see Figure 1). We can use this theory to explain why there is an empirical law linking the pressure,

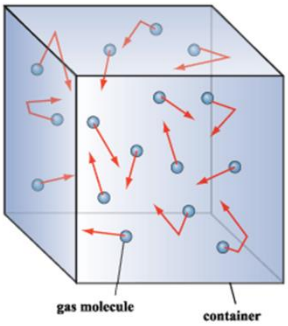

Fig. 1. The kinetic gas theory

the volume and the temperature of a gas, the so-called law of Boyle-Gay Lussac or ideal-gas law (pV/T=constant) which received its final form in the early nineteenth century, after preparatory work going back to the seventeenth century. Now this theory represents all gas particles as identical, even when the gas is common air, which we know to be a mixture of several gases – we have empirical methods for separating them. It is not plausible that the particles from which each of these different gases are composed are truly identical and that any collision between any two particles is always perfectly elastic. Further, if the gas consists of small particles, so presumably do the walls of the container. These walls, however, are not so described: they are described as perfectly flat structureless surfaces against which the gas particles bounce off completely elastically. This is hardly a description we can reasonably expect to be descriptively true. But both the explanatory and the predictive work of theories is done through deductive reasoning, and so we must assume the statements of the theory to be strictly true, because deductive reasoning knows only truth and falsity, that is, strict truth and strict falsity. In particular there seems to be no room for aspects that, intuitively, are part of our interpretation of how theories describe the world, such as idealization and approximation.

Let us therefore try to develop a view that leaves room for these aspects. For simplicity's sake let's remain with physics and take the simple pendulum. A physical treatment of the pendulum, as an application of classical mechanics, starts with a representation of a pendulum *from the perspective of physical theory* (see Figure 2).

The generality of a theory like classical physics, or of any theory for that matter, means that its statements are, logically speaking, conditional: *if* there are particles having masses and positions in space that can change, and *if* there are forces working on them, then the relation between the masses, accelerations and forces is given by $F=ma$. To see what the theory implies for a concrete case, that case must be given first, in a format that matches the concepts the theory deals in. And nothing that the theory does *not* deal in needs to be given, since we would not know what to do with that part. The picture of the pendulum in Figure 2 does exactly that: it tells us what movement the rod can make and it tells us how to characterize this movement through an angle θ with the vertical which can have a variety of values. It also tells us what the lengths of the moving parts are, and, although not explicitly indicated, it tells

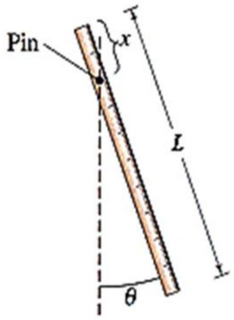

Fig. 2. The simple pendulum

us that the rod has a particular mass and that a particular force, the force of gravity, works on it. For this situation, we can derive from the theory of classical mechanics that the rod will enter a periodic motion of swing first to one side, then to the other, and that the period T of this motion is given by the well-known formula $T=2\pi\sqrt{(L-x)/g}$. (Note that this formula is not exact according to classical mechanics itself; it assumes that θ is small, because if we don't do this we are unable, even in this simple case, to derive a mathematically exact expression for T as a function of the other variables.) With this formula we can, for any particular pendulum, by entering the values of L, x and g in the formula, calculate the period T of the swing it would get into when let go from any non-vertical initial state.

The picture, together with the parameters x, L, θ and m (the mass) and F_g (the force of gravity), do not describe any actual pendulum in the world. Even if we give these parameters precise values that we measure on some rod that we hang on some pin so that it can swing from that point, the picture and the parameters do not truthfully describe that pendulum. An infinite number of properties that the rod-pin system – the pendulum – has are not present in the description. For some of these – the rod's colour, for example – the theory we have available – that is, the sentences that state the law-like relations that exist between these parameters and that we use to predict their values as they change in the course of time, is entirely blind. For others, this is definitely not so. The picture mentions only one force, whereas we have reason to believe that more forces work upon the rod: the friction between rod and pin and the friction between rod and surrounding air, to mention the two most obvious ones. If these were added, the prediction derived from the theory would certainly change.

Let us take stock of the various ingredients of the situation ad their precise relationships, both mutually and with the real world. First, there is an arrangement of symbolic entities, partly pictorial and partly linguistic (but it could be exclusively pictorial or exclusively linguistic). Insofar as this arrangement of symbolic entities truthfully describes anything, it described a pendulum existing in some conceptual, mental or imaginary world, a world that that can even be imagined to contain nothing but this pendulum and the force of gravitation. The arrangement of symbolic entities can be termed the description of this conceptual or imaginary world, or of the part of this world that contains the pendulum. As a description of that object it is by definition true. To express this, we call it a *specification* of the world, or the pendulum object existing in this world. The theory, finally – the arrangement of linguistic entities that gives some dynamic laws to which the properties that the conceptual object has are subject – is also by definition strictly true in the conceptual world, and therefore true for the pendulum object. That is, after all, how the conceptual or imaginary world is conceived, came into conceptual being so to speak.

The three things – conceptual object or world, its specification, and a theory valid in that world or for that object – together form a *model* of or for any real pendulum, say, some rod that we suspend from a single swing point. It should be noted, however, that in practice the word model is used very sloppily. Ignoring the many cases where the word 'model' is used for something else entirely, such as a method, an approach, and so forth, it is often used so as to indicate just one or two of the three ingredients distinguished above. Especially in physics, we are often dealing with theories that are much more general than any particular model in which they are applied. The theory is then typically not considered to be a part of the model. This applies to the pendulum, but it also applies to the kinetic gas model, where the model's specification states that the collisions between gas particles are completely elastic, but where the laws of mechanics that state what happens in completely elastic collisions (conservation of momentum and of kinetic energy) are not considered to be part of the model itself, because they have a much wider validity. In other cases, for example the Lotka-Volterra model, which describes the interdependent developments of the sizes of a population of predator animals and a population of prey animals (see Figure 3), the equations are definitely considered to be part of the model and are even seen as the core of the model.

x = population size of prey animal

y = population size of predator animal

$$^{dx}/_{dt} = \alpha x - \beta xy$$
$$^{dy}/_{dt} = \gamma xy - \delta y$$

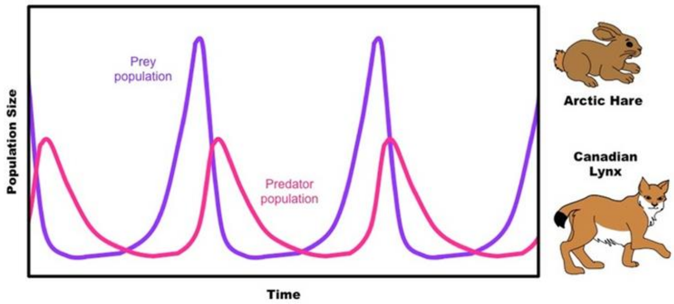

Fig. 3. The Lotka-Volterra model

Apart from whether the theory is seen as included in the model, there is a difference between researchers in whether the conceptual world or object is explicitly seen as a separate component of any model. Some researchers think a model is exhausted by the linguistic and/or pictorial ingredient and there is no need to assume there is actually something of which that part is a description or specification, whereas others emphasize that conceptual or imaginary component as the actual model, and accordingly restrict the use of the term to just that element. The latter view allows us to maintain a single interpretation of language, or at least of its declarative part, that is, ignoring questions and imperatives, and to maintain as well its relation to logic as ordinary language purified. Any declarative sentence must describe some situation, and that description is either true or false. The opposing view is motivated by the feeling that there should be no role in science for something so subjective and ungraspable as a conceptual or imaginary world or object, or that our interpretation of what scientific theories are should not be made dependent on such a notion.

If we ignore the conceptual or imaginary object, it becomes difficult to clarify the relation between the pictorial-plus-linguistic component of the model and the real world. Just as that part is by definition true of a conceptual model, it is by definition false of the real world. This severely reduces our possibilities of linking the model description to the real world. The notion of a conceptual or imaginary world or object serves exactly the function of mediating between the pictorial and/or linguistic part and the real world. It is related to the linguistic part in the same way as ordinary language is related to the ordinary, empirical world, namely is describing it, either truly or falsely, where we are able to control for the absence of falsity in the case of models. And it is related to the real world in a way that corresponds to the sort of comparison that is at issue there: not a relation of true of false but a relation of similarity or approximation. The spatial connotation that the word 'approximation' has is entirely apt here: the relation of a conceptual model to the real world can be seen as being at some distance from the real world. The larger the distance, the less is the model 'like' the real world. So although every model is necessarily false, we suppose that one model can be, we are inclined to say, 'less false' than another. It is exactly because the 'true' and 'false' do not allow of any ‘more’ or ‘less’ that we have to look at the relation between model and world differently, as one of more or less resemblance or more or less distance.

The distance from the real world at which a model is set is brought about by two distinct activities: *abstraction* and *idealization*. Abstraction is an impoverishment first of all of entities, things. Many objects that exist in the real world are not contained in any model used to describe part of that world and make predictions concerning it. This includes things that must exist in the real world in order for those things that are shared with the model to exist. The wall, for example, or the stand, as the case may be, to which the pin on which a pendulum swings is attached, is not contained in the model; there it exists as it were free-floating. In the model, beyond the pendulum nothing exists except the force of gravitation, and in the case of the kinetic gas theory nothing exists beyond the container and the gas it contains, not even the force of gravitation. In the Lotka-Volterra model of a predator and a prey population the food that the prey animals eat is not contained, nor is the soil on which both populations live.

We also speak of abstraction if entire classes of properties that things have are not contained in a model. The pendulum model abstracts from colour, for example, or material from which the rod is made. Nothing in the model pendulum world has colour or is made of anything. In the case of idealization, in contrast, a thing or property class is also present in the model world but is distorted in being simplified. The force of gravity is, strictly speaking, slightly weaker near the top of the rod than it is near the lower end of the rod, and may also vary somewhat when we move in the horizontal direction. But in the model the force of gravity is assumed to be uniform. The molecules from which a real sample of air is composed collide, very likely, not completely elastically, nor do they collide completely elastically with the walls of the container. The range of collision forms existing in reality is simplified to one uniform type in the model. In practice the terms ‘abstraction’ and ‘idealization’ are not used so that it always matches the distinction made here. The ignoring of friction in the pendulum model, for example, will generally be called an idealization, whereas according to the way the distinction is introduced here it is actually an abstraction. And even when this way of distinguishing between abstraction and idealization is accepted, the precise borderline separating the two is difficult to draw.

Intuitively the distance between a model and reality can be reduced by reducing the idealization and the abstraction. This ‘can’ is, however, ambiguous. Especially in physics, the background theories are in principle able to deal with lowering the degree of idealization, for instance by introducing less-than-completely elastic collision in the kinetic-gas model or friction in the pendulum model. In other cases, however, the relevant theory is completely mute as soon as we depart from a particular idealization. This is of course especially true in the case where a theory is unique to a model. To make the model less idealized would then require a new model, with an extended theory. Similarly, to make a model less abstract will not lead to any gains if a theory that is able to describe the behaviour of the newly admitted entities or newly admitted property class is lacking.

Talk of models that approximately describe reality in the sense explained above – that is, the models are completely specified mental or imaginary worlds or objects that resemble reality or approximate reality, or that picture reality as being like that – makes sense only to a someone who has a realist perspective on theories: someone who believes that one of the goals of science is to explain observable phenomena by revealing the underlying structure of reality, and the underlying mechanism at work on this structure, through which these phenomena are brought about. It does not make sense to someone with an instrumentalist perspective on theories: someone who believes that theories are nothing but instruments to predict observable phenomena. Then how can a realist and an instrumentalist scientist communicate? The answer lies in the fact that realist scientists may have their own beliefs but must nevertheless work with theories and models in the same way as the instrumentalist. The extent to which a model resembles or approximates reality cannot be established directly, if at all. The validation of a theory or a model must proceed through its observable consequences. In order to do that, the theory or model must be connected to the world of empirical phenomena. In the case of the kinetic gas theory, for example, it must be investigated how the velocities and collisions of gas particles and their further effects make themselves felt as the pressure and temperature of a gas. Or at least, that is how a realist scientist would phrase it. An instrumentalist scientist would say that it must be investigated how from the formulas of the classical mechanics made up of x_i-s, m_iv_i-s and $½m_iv_i^2$-s, *as if* these were the positions, momenta and kinetic energies of small particles, statements can be derived concerning the pressure and temperature of a gas. Both scientists need the same sort of thing: *correspondence rules* linking terms from the theoretical language, the language in terms of which the theory is stated, to terms from the observational language, which is used to express the relations between empirical concepts.

It is important to realize that such correspondence rules can be derived neither from the empirical data nor from the theory. Both parts are conceptually isolated from each other. This is obviously so for the instrumentalist, for whom the theoretical language contains nothing but abstract cogs and wheels whose only role is to serve in the operation of a conceptual machine for prediction but which do not refer to anything. For the realist the theoretical and observational languages may seem to share certain terms, but these terms do not for that reason stand for the same concepts. The observational term 'temperature' as operationalized through thermometer readings is not identical to the term 'temperature' as it occurs in the theory of thermodynamics, as can already be concluded from the fact that the former is measured on an interval scale (degrees Celsius, for example, or degrees Fahrenheit) but the latter on a ratio scale (Kelvin). Likewise, with regard to the kinetic gas theory, 'volume' in the observational language is defined by the dimensions of the container holding the gas, whereas 'volume' in the theory is defined as the segment of Newtonian space that the gas particles have available for their motion. It is a mere assumption that these two come down to the same thing, and we can think of many ways in which they would not come down to the same thing.

One of these led to an important correction of the ideal-gas law by Johannes van der Waals in 1873, when he pointed out that if the gas particles are not point-like but have some finite size, the space available for their motion is smaller than the observational volume defined by the container's dimensions, due to the fact that all particles together occupy a significant part of that latter space which is then not available for the motion of gas particles. This allows for a revised model closer to reality, which takes this (and also other refinements) into account and which predicts – on the assumption that there are no further things at work causing additional separation between observational and theoretical volume, and on the assumption that other correspondences between theory and observational level are maintained – a more accurate joint relation of observational pressure and temperature to observational volume, the Van der Waals equation. In general, correspondence relations will be chosen so that a maximum number of relations manifested at the empirical level follow, such as the ideal-gas law or the relation between period and length of a pendulum or the undulating development, separated by some phase difference, of the sizes of a population of predators and a population of prey animals, which are then open for confirmation or disconfirmation. The activity of (dis)confirming these empirical consequences of a model – more precisely of a model-theory-correspondence rules complex – is called its *validation*. A model and/or theory survives as long as these relations are confirmed and no competing theory and/or model emerges that furnishes an even larger number of confirmed empirical relations.

Figure 4 portrays the complex relations existing between the various parts of on the one hand the linguistic apparatus we use when we do science and on the other hand the aspects of the world our science is directed at: first the phenomena that directly affect our sensory apparatus and in the acceptance of which as real all scientists and engineers are united; second the observational language serving to describe these phenomena – both the ones we collect in the wild through careful and methodical observation and the ones we bring forward in laboratories through experimentation, in short, the totality of our evidence; third the theoretical language, which may either be seen as a purely abstract machinery for reasoning and calculation or as describing abstracted and idealized fragments of some world; and fourth and finally the perhaps infinite deepness of the world as it extends beyond the phenomena, of which realist scientists hope that the world segments that they take our theories and models to describe are either a faint echo or a mirror image. To the left is depicted, for the case of the kinetic gas theory, the minimal relation that all scientists, be they instrumentalists or realists, share: the worth of any theory or model can only be established by deriving consequences at the empirical level from the conjunction of theoretical apparatus and correspondence rules. To the right is depicted, for the same case, the possibility of interpreting the theoretical apparatus as a model that describes some abstracted and idealized world or semi-world, which approximates the world as it really is, beyond the part that is given to us in perception.

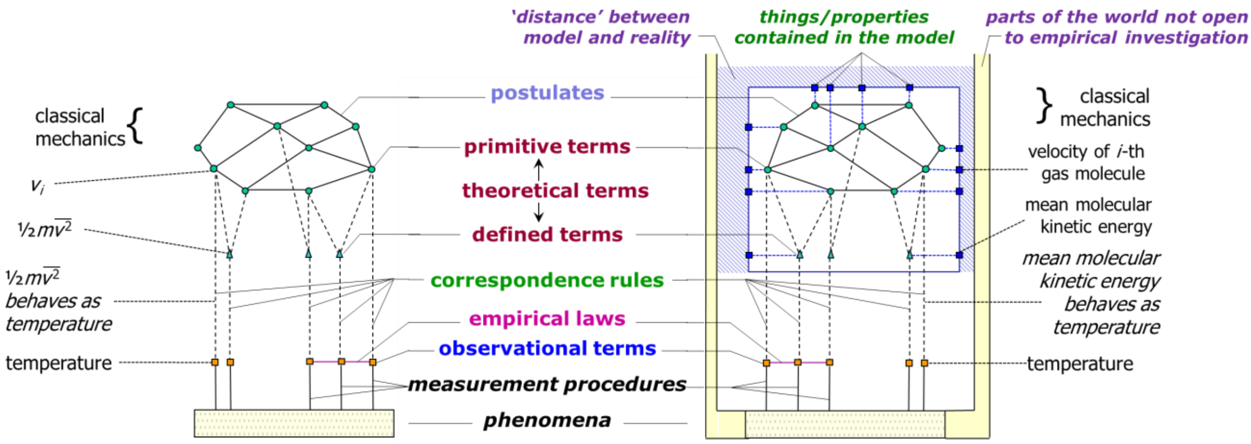

Fig. 4. The relation between the empirical phenomena, the observational language of science, the theoretical language of science, the specified model, and the structure of reality beyond the phenomena.