mot1442 – Text 2 (version without logic)
The validation of scientific claims

1. *Prediction and the role of laws and theories in science and technology*

Of primary importance in the practices of science and technology is the notion of a scientific law. Traditionally science as a whole has served the lofty purpose of making us ‘understand’ the universe we inhabit, and it has done so by revealing the structuredness and regularity at work behind the enormous variety of natural phenomena. Of a much more practical importance, we are interested in the ability to make predictions concerning these phenomena, in order to increase our chances of individual and collective survival and to make our lives more comfortable and rewarding. Accurate prediction depends crucially on laws. We may distinguish stronger and weaker forms of prediction. In the latter case, we may have reason to expect a certain state of affairs to obtain, or a certain phenomenon to occur, but we cannot be sure it will. Generally, however, when we apply scientific knowledge in a technological setting, we go for predictability in the strongest sense. If a beam is placed on some supports and a load is to be placed on top of it, we are interested in the *truth*, and nothing less, of the statement that the beam will support a load of 500 kilograms. We can be convinced of the truth of such predictive statements if we succeed in deriving them as conclusions of valid arguments with true premises. And it is among the premises of such arguments that scientific laws do their work.

Take the following example of a prediction that a particular object made of metal will conduct electricity:

All metal objects conduct electricity
This object is made of metal

Therefore, it will conduct electricity

This argument is indeed valid, or more precisely, an instance of a valid reasoning scheme. Its validity can be seen more directly by recognizing that the first premise is logically equivalent to ‘No object made of metal fails to conduct electricity’. On the basis of this argument, we can be convinced that any particular object will conduct electricity, provided that it is made of metal and provided that it is true that all metal objects conduct electricity, i.e. that this is a law of nature.

2. *Establishing the truth of generalizations: enumerative induction*

This example shows that accurate prediction, where we make a claim concerning a future state of affairs that we straightforwardly expect to be true, depends crucially on the truth of the sort of general statements that make up the laws of science. The extent to which we can be convinced of the truth of our predictions depends, therefore, on the extent to which we can be convinced of the truth of these laws, or of the theories of which they form a part. And that brings us to the question how we can become convinced, in turn, of *their* truth. Since scientific research and of a considerable portion of technological research are expressly aimed at the ‘uncovering’ of these laws, this question becomes the question of the justification of the method of science. This questions turns out to be a much more difficult one to answer that the question how science is able to make accurate predictions.

Naively, it could be thought that such laws are 'based on the facts', either the facts of nature that any person can obtain by accurate observation, or by facts we create ourselves in the form of experiments. For the example at hand, for instance, we could accumulate a large number of objects that we have ascertained are made of metal and that we test for conductivity. Suppose we tested *n* of these metal objects, a_1, …, a_n, and all of them would test positively. Our problem how to arrive at the truth of scientific laws would be solved if we could reason as follows:

Object a_1 is made of metal and conducts electricity
Object a_2 is made of metal and conducts electricity
…
Object a_n is made of metal and conducts electricity

All metal objects conduct electricity

This cannot be the whole story, of course, since the conclusion is asymmetric whereas all premises are symmetric. This can be amended by adding a premise that represents our background knowledge that the relation between 'being of metal' and 'conducting electricity' is asymmetric: conductivity is an attribute that also things made of other substances than metal can have and occasionally indeed have. This blocks 'All metal objects conduct electricity' as a conclusion.

Let it be assumed, then, that the premises 'Object a_1 is made of metal and conducts electricity', 'Object a_2 is made of metal and conducts electricity', ..., 'Object a_n is made of metal and conducts electricity' correctly enumerate our evidence. But the truth of these premises, however large *n* is, does not guarantee the truth of the conclusion. It remains possible that all premises, i.e. the totality of the evidence at the moment of evaluation, are true, whereas the conclusion is false: some hitherto unexamined piece of metal may, on inspection, prove to lack the property of conductivity.

An argument of the above structure is known as an argument by induction, or more precisely (in view of things that are said below on inductive reasoning as a general concept) an argument by *enumerative induction*. A weaker form is where we conclude not to a generalization, i.e. a prediction that all hitherto unexamined metal objects in fact conduct electricity but to a concrete prediction that some arbitrary but particular metal object is conductive:

Object a_1 is made of metal and conducts electricity
Object a_2 is made of metal and conducts electricity
…
Object a_n is made of metal and conducts electricity

If the next object I observe is made of metal, it conducts electricity

This argument can be seen as a combination of the argument by enumerative induction and the prediction that we started this chapter with. It is, of course, equally invalid. No finite amount of evidence showing a perfect correlation between instances of the property of being made of metal and the property of conducting electricity – and all our evidence is necessarily finite – will guarantee that we will continue to observe this perfect correlation in the next item under examination.

This negative outcome is closely related to the character of natural laws. Not all generalizations, that is, all statements of the form ‘Everything that is A is also B’, where A and B are arbitrary properties, are scientific laws. The statement ‘All plants in this garden are in bloom’ is an example of a statement of this form that is not a law of nature. The point is not that this statement is restricted to a particular place or time. Many bona fide laws of nature are similarly restricted. ‘All freely falling objects fall with the same acceleration of 9.8 m/s^2’ or ‘All fossils of mammals are less than 200 million years old’ are examples of generalizations that are true only when restricted to the Earth, but they are no less law-like for that. What distinguishes natural laws from *accidental* generalizations like ‘All plants in this garden are in bloom’ is that the former are open-ended: they cover an indefinite number of cases. If we pass the same garden a week later, we could find that several of the plants are no longer blooming. In contrast, we can keep on dropping objects from a height and they will still fall with an acceleration of 9.8 m/s^2, and we can continue digging up fossils of mammals and they will still prove to be younger than 200 million years. The number of cases covered by a natural law is potentially infinite, although it need not actually be infinite: the number of mammals that have actually lived on earth is necessarily finite, and so is the number of falling objects on earth, given the earth will be destroyed by the expanding sun in about 5 billion years. Another way of expressing the fact that natural laws cover an indefinite number of cases is that natural laws support counterfactual statements. We can say, of an object that we hold fixed and that never in its lifetime will actually fall freely, that if we were to let go of it, it would fall with a constant acceleration of 9.8 m/s^2. In contrast, if we observe that all plants in a particular garden are in bloom, it is not true of another plant growing somewhere else that, if it were a plant in this garden, it would be in bloom.

This failure has led to proposals for arguments by enumerative induction such that they become true. For example, by adding extra premises saying that ‘the future is like the past’ or that ‘nature is uniform and true to itself’. This would be to beg the question, of course. The truth of the conclusion would follow provided this extra premise is true. But how could we become convinced of its truth? Saying that this extra premise is true is just another way of saying that we are justified in concluding that all metal objects conduct electricity if all hitherto examined metal objects have proved to conduct electricity. But the justification of this way of reasoning is precisely what is at issue, and we have not advanced an inch beyond our first try that resulted in a failure: the conclusion does not follow deductively through a valid logical argument.

This difficulty, often referred to broadly as the problem of induction, has been discussed with much intensity ever since the philosopher David Hume brought it forward in the first half of the eighteenth century. We cannot say that Hume discovered the weakness of the link between evidence and our predictions concerning the future. In fact Francis Bacon, who is often portrayed as the inventor of the inductive method, more than a century earlier, knew very well that extensions of our experience into the future carry no certainty. When we reason from our experiential past to our not-yet-experienced future, we do something different from what we do when we reason deductively, as we do when we derive a prediction from a law and a condition. The problem of induction is the question whether what we do when we argue inductively is justified, or to what extent it is justified. It might be thought that we have just shown that it is not justified. However, that attitude presumes that the only way to justify a particular piece of reasoning is to reconstruct it as a deductively valid argument, and that might be much too narrow. Perhaps an argument can be non-deductive and still justified.

3. *Establishing the truth of generalizations: the hypothetico-deductive method*

This possibility, however, will only be discussed in detail in another chapter. The possibilities of showing scientific reasoning toward generalizations to be a form of deductive logic are not yet exhausted. Another attempt presents scientific reasoning as a form of *hypothetico-deductive* reasoning. The argument by enumerative induction mimics as it were our psychological process of 'discovering' a scientific law by deriving it as the conclusion of an argument. That this is the way we arrive at the universal statements of science may look plausible for low-level empirical generalizations like 'all metals conduct electricity', but becomes less so for high-level theories like classical mechanics or classical electrodynamics, or Darwin's theory of evolution by natural selection. Such theories are not separated from the experimental evidence by just the single step of inductive generalization. Compare the following two arguments:

i. Object a_1 falls freely and undergoes a constant acceleration of magnitude c
Object a_2 falls freely and undergoes a constant acceleration of magnitude c
…
Object a_n falls freely and undergoes a constant acceleration of magnitude c
___
All freely falling objects undergo a constant acceleration of magnitude c

ii. Object a_1 falls freely and undergoes a constant acceleration of magnitude c
Object a_2 falls freely and undergoes a constant acceleration of magnitude c
…
Object a_n falls freely and undergoes a constant acceleration of magnitude c
___
Newton first, second and third law of dynamics as well as Newton's law of universal gravitation hold

Although both arguments are invalid, the first one presents itself as at least a prima facie plausible candidate for a convincing argument. The second one fails this entirely. Regardless of the validity problem, the method of enumerative induction therefore seems to leave a large number of the universal statements of science unaccounted for.

The hypothetico-deductive method is specifically aimed to handle not just low-level empirical generalizations but also high-level theories. In this method, the universal statements at issue are introduced explicitly from the start, rather than derived as the conclusion of an argument. However, they are proposed as *hypotheses*, statements about the truth value of which we are, as yet, uncertain. This explains the 'hypothetico' in hypothetico-deductive method. In the next step, a prediction of a particular phenomenon is derived from the hypothesized statement, by logical deduction, as was shown above. For example, from Galileo's law, 'All freely falling objects undergo a constant acceleration of magnitude c', it can be derived that if we create a certain condition, i.e. we let object a fall freely, then a will be observed to fall with constant acceleration of magnitude c. But we can derive the same prediction, object a will be observed to undergo a constant acceleration of magnitude c if we let it fall freely, also from the conjunction of Newton first, second and third law of dynamics plus Newton's law of universal gravitation. This second step is clearly deductive. Whether it alone explains the 'deductive' in hypothetico-deductive method is not so clear, however. We are still nowhere near a judgement on the status of the universal statement. Another step is needed. In this final step, the prediction is compared with the empirical facts, and, hopefully, we are allowed to draw a conclusion on the basis of the comparison. Let T indicate an arbitrary theory under scrutiny, let C be some empirical condition that we can either create during an experiment or can observe to be the case somewhere in nature, and let E be the empirical fact that is predicted to be the case in condition C if we assume the hypothesized theory or law T to be true. So we have that if theory T is true, then if condition C obtains, we can observe E to be the case. Now two things can happen: either condition C obtains and E is indeed observed, or condition C obtains and something different from E is observed. In the former case, the conclusion we would like to draw is that the hypothesis is true, in the latter case that the hypothesis is false. But can we?

If we observe something different from E in condition C, we can indeed draw the conclusion that H is false, by way of a deductively valid argument. Whatever it is that is observed, if it is different from E, we know that the following claim is true: C obtains but E is not the case. Indeed,

If theory T is true, then if condition C obtains, we can observe E to be the case
C obtains but E is not the case

Theory T is false

is a valid argument. And it is also a sound argument: the prediction that in condition C, E is the case follows itself by deductive logic from T, which, after all, is a statement, be it a very long and highly structured one. Therefore we can be confident that the first premise of the argument is true. And the second premise we know to be true by observation. So, on the basis of the evidence 'C obtains but E is not the case', the hypothesis has been *falsified*.

What if condition C obtains and E is indeed observed to be the case? If a conclusion could be drawn here, we would have the contrary case of falsification, namely verification. Reviewing the arguments for falsification, it can be easily seen that in this case the argument resulting in the conclusion that T is true is *not* deductively valid. If we were to take this as valid, we would be operating as if we can turn our first premise, 'If theory T is true, then if condition C obtains, we can observe E to be the case' around, so that we get 'If condition C obtains and we observe E to be the case, then theory T is true'. Logic tells us we cannot. Return to our simple hypothesis that all metals are silver-coloured. Can we conclude that this is true once we have observed a sample of chrome to indeed be silver-coloured? Clearly not. To accept

Object *a* is made of metal and is silver-coloured

Every object made of metal is silver-coloured

as a valid argument would amount to the weakest form of enumerative induction: generalizing from a single observation. Notice that to propose 'Every object made of metal is silver-coloured' as a hypothesis, we need not yet have observed any samples of metal. we could have arrived at this claim just because it occurred to us that silver would be a nice colour for metals!

So the hypothetico-deductive method only partly delivers what we hoped for. It can establish the falsity of a universal statement by a deductively valid argument, but not the truth of such a statement. This would not be a problem if we could arrive at the true theory by elimination of false theories. Very often in the history of science, two competing theories existed alongside each other during some period. Examples are the Ptolemaic and Copernican views of the universe, classical Newtonian mechanics and Einstein's theory of relativity, the particle theory of light and the wave theory of light, the phlogiston theory of combustion and the oxygen theory of combustion, and so forth. A test for a condition where two or more theories predict different things to happen is called a *crucial experiment*, since the prediction of at most one theory can be borne out. This suggests that our doubts concerning which theory to accept will be resolved by the experiment or observation. Indeed, if we somehow could become convinced that of two competing theories, one must be true, we would be able to arrive at its truth by the deductively valid falsification of its competitor. For two theories, let's call them T_1 and T_2, we can express the idea that one of them is true by 'Theory T_1 is true if and only if theory T_2 is false' Now suppose that T_1 is falsified by the evidence, which means that we have 'If theory T_1 is true, then if condition C obtains, we can observe E to be the case' and additionally 'C obtains but E is not the case' (because some evidence E* other than E is observed to be the case). Now

Theory T_1 is true if and only if theory T_2 is false
If theory T_1 is true, then if condition C obtains, we can observe E to be the case
C obtains but E is not the case

Theory T_2 is true

is a valid argument. We would have established the truth of T_2 by elimination of its only competitor T_1. This will not work, however, because we can never become rationally convinced that one out of a given set of theories must be the correct theory about a particular class of phenomena. Although this is not a thing that we know how to prove, it is highly implausible that we are able to come up with all the logical possibilities of how to account for a particular domain of phenomena. And that we would be able to do this is necessary for verification of one theory by elimination through falsification of all competing theories to succeed.

Some people contest that this is not so for our theory for the totality of all phenomena, or – which to these people is easily the same thing – for the basic building blocks of the world. They hold that we are able to phrase all the logical possibilities, or even the unique theory that can make sense of them. But this position seems to assume that we can have access to the totality of all phenomena, or that the currently recognized basic constituents of the world, elementary particles or strings, are exactly these and not complex things made up of more elementary things after all. And how can we be sure of that?

4. *The poverty of falsificationism; the ubiquity of assumptions*

Still, it has been forcefully argued, notably by the philosopher Karl Popper, that falsification is all you need in science. According to Popper, science proceeds exclusively through the, deductively valid, elimination of false theories. Verification of true theories in a deductively valid way is impossible, but this does not compromise science: science has no need for verification. This view entails a radical modification of our conception of scientific knowledge: what we know about the world is what the world is *not* like. All our work does not enable us in the least to say what the world *is* like. A particular theory may see its predictions turn out true again and again, but this does not give it a higher status with respect to other, less well corroborated (Popper's favourite term) theories. It could be proved false by the next piece of evidence just as easily. If falsification is all the method we have in science, there are no grounds for favouring a theory that has survived many test over a theory that is new to the scene.

This may be an acceptable position for someone whose interest is purely contemplative, involving a sceptical attitude that values the avoidance of believing anything false over all other considerations. It seems unacceptable, however, to anyone whose interest in scientific knowledge is, at least partly, practical, directed at applying knowledge. An engineer who undertakes to build a bridge or design a new jet aeroplane is not likely to accept that there is nothing to choose between alternative theories available to base her models or calculations on, on the argument that the only sensible thing we can say about these theories is that they may all be false, although they have not yet been proven false. In a philosophical mood, such an engineer may be prepared to admit that the truth of none of these theories can be guaranteed, but surely there are better and worse ones to pick?

Additionally, this position is difficult to square with the historical record of science. It may not be such a difficult conclusion to stomach with respect to the low-level empirical generalizations that presentations of the problem of induction usually focus on. However many samples of metal we show to conduct electricity, can we really be sure that *all* metals conduct electricity? The structure of modern science is such that the amount of trust we have in the truth of this generalization depends more on how we can locate it in the network of our high-level theories than in the number of samples we have observed. We have a theory about metals, about the microstructure that is common to all metals, from which it follows that no substance can have this microstructure and not conduct electricity. Our acceptance of empirical laws of the sort that all metals conduct electricity therefore depend just as much, if not more, on the grounds we have for accepting the underlying theories than on the experimental results that test it directly.

The inability to say anything distinctly *positive* about theories that are borne out by their predictions seems therefore an intolerable weakness of the hypothetico-deductive method, in its Popperian form. But the method's ability to be distinctively *negative* does not stand either, on closer scrutiny. This is because the derivation of predictions from theories is hardly ever a straightforward matter. It comes close in the case of low-level empirical laws like 'All metals are silver-coloured', though even here not quite close enough. It seems that once we see that copper is reddish-yellow, this law is out. If, that is, we accept that copper is a metal. Since the law only makes sense if we know how to identify metals, usually there will be criteria in place that decide whether copper is a metal or not. However, we could always change the criteria, and deny that copper is a metal. To be sure, there is a price to be paid for such *ad hoc* strategies to save a law from falsification. Many of our other laws would turn out a bit awkward, as awkward as, for example, 'All metals and in addition copper conduct electricity'. As soon as we start to develop theories that explain what is common to all metals in terms of their microstructure, it becomes increasingly strained to maintain that copper is not a metal. Therefore, one hardly ever, if at all, sees such *ad hoc* rescue strategies employed in science.

But it is always a possibility, there is no rule of logic that blocks it. Strictly speaking this does not indicate a weakness of the hypothetico-deductive method itself but in its scope: one of the premises in the logical argument of falsification is retracted and we no longer have an argument.

Now take a slightly more complicated case, Galileo's law of free fall. A famous story in the mythology of science relates how Galileo demonstrated this by dropping two objects of different weight simultaneously from the leaning tower of Pisa, demonstrating that they hit the ground at the same time. One of the reasons for thinking that this story is a myth is that it is well-documented that another Italian scholar, Giorgio Coresio, actually did drop two objects of different weight from the tower in 1610 simultaneously and showed that they did *not* hit the ground at the same time, although the difference was smaller than most Aristotelians of the day may have expected. So Galileo's law is out. But Galileo's law is a consequence of the theory of Newtonian mechanics in combination with the law of universal gravitation, although Galileo did not yet know this, of course. With the notation used already above, we have, *prima facie* at least, 'If the combined theory N of Newtonian mechanics and universal gravitation is true, then if objects *a* and *b* fall freely over the same distance while having started at the same moment, they will hit the ground at the same time'. Coresio's observed 'Objects *a* and *b* fell freely over the same distance while having started at the same moment but did not hit the ground at the same time'. We must conclude that Coresio's observation falsifies N, the conjunction of Newtonian mechanics and law of gravitation. We are convinced, however, that if people were to conclude from this experiment that N is false, they would make a mistake. So the whole argument has to go, since otherwise we are logically compelled to accept the falsity of N. (We are ignoring the theory of relativity, which has shown us that the Newtonian conjunction N is indeed not true, but this experiment is not the reason why we consider it false and irrespective of its falsity we judge that during the entire period from 1600 to 1900, prior to the articulation of theory of relativity, someone who concluded to the falsity of N on the basis of Coresio's experiment, or a repetition of it, would be making a mistake.

To show that the entire argument is faulty is not difficult. The discrepancy between theory and prediction is due to the friction of the air affecting the two objects differently. But no statement mentioning air friction figures in the argument. Indeed, once we give the matter some more thought, we should realize that the prediction 'If objects *a* and *b* fall freely over the same distance while having started at the same moment, they will hit the ground at the same time' follows from N only if we add assumptions on how the particles composing the surrounding air interfere with the fall of the objects *a* and *b*. After all, if the air particles exert forces on a falling object, its motion is determined by the sum of all forces, resulting in an accelerated motion that is likely to depend on the object's form and weight. The prediction follows only if the assumption is made that the presence of air does not affect the rate of fall of free-falling objects, or, more accurately, that insofar as air affects this rate it does so to the same extent for all objects. The relation between theory and prediction has, therefore, not been represented accurately; the correct relation is 'If the combined theory of Newtonian mechanics and universal gravitation is true and if the presence of air does not affect the rate of fall of free-falling objects, then if objects *a* and *b* fall freely over the same distance while having started at the same moment, they will hit the ground at the same time', where A stands for the additional assumption that. The structure of this statement is 'If theory N is true and additionally assumption A is true, then if condition C obtains, effect E can be observed', where in this case A corresponds to 'The presence of air does not affect the rate of fall of free-falling objects', C to ‘Objects *a* and *b* fall freely over the same distance while having started at the same moment’ and E to ‘Objects *a* and *b* hit the ground at the same time’. However,

> If theory N is true and additionally assumption A is true, then if condition C obtains, effect E can be observed
> Condition C obtains but E is not observed
> ―――――――――――――――――――――――――――――――――
> Theory N is false

no longer is a deductively valid argument. Instead, the valid argument goes like:

> If theory N is true and additionally assumption A is true, then if condition C obtains, effect E can be observed
> Condition C obtains but E is not observed
> ―――――――――――――――――――――――――――――――――
> Theory N is false or assumption A is false

We are no longer forced to accept the falsity of N, but only to accept the falsity of either N or A (and possibly both). And the correct step, we are convinced, would be to accept the falsity of A but not N: out in the open, the surrounding air *does* affect the rate of fall of falling objects, and it does so to a different extent in objects of different weight.

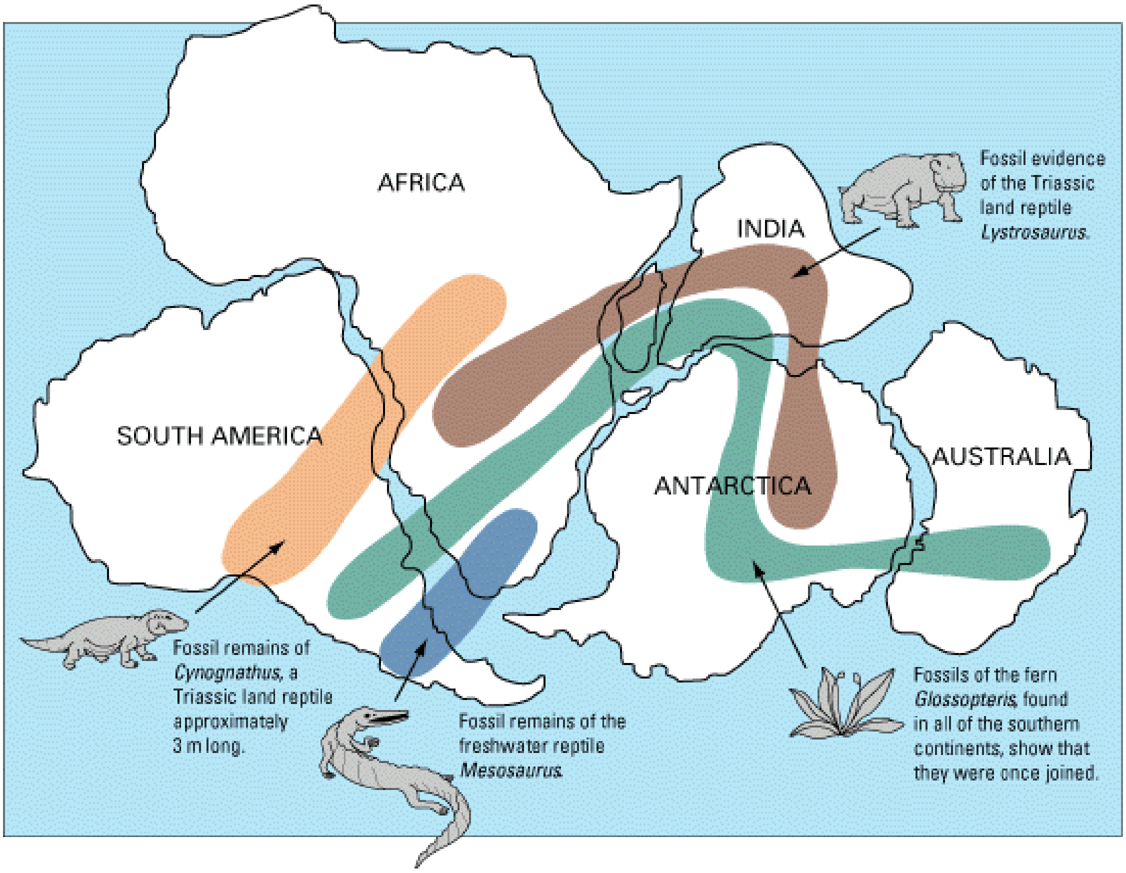
So we have found another fundamental weakness in the hypothetico-deductive method. For any theory numerous assumptions are required to predict what will be the case in a particular empirical condition. Or, alternatively, any theory can be further analysed into a core, which contains the actual explanatory framework, and additional assumptions, which could be replaced by other assumptions while the core theory remains intact. But this means that if the prediction is not borne out, what we can maximally conclude is that either the (core) theory is false or at least one of the assumptions is false. There is nothing that allows us to pinpoint the difficulty more accurately, to shift it either in the direction of the theory or in the direction of the assumptions. This means that the hypothetico-deductive method is as good as powerless to settle the life or death of theories in real-life science.

Case study: Continental drift

Already in the seventeenth century people noticed that the coastlines of Africa and South America fitted together surprisingly well. This led to occasional speculation that these continents had been connected in the past. But in 1915 the issue left the realm of speculation. In this year the German geologist Alfred Wegener published his book *Die Entstehung der Kontinente und Ozeane* (The origin of the continents and oceans), presenting a truly revolutionary theory, which held that all continents have wandered all over the surface of the earth during our planet’s geological past and are actually still doing so. During a certain period all current continents even formed one ‘supercontinent’, in which not only South America lay directly against Africa but also North America directly against Europe, while against the current eastern coastline of Africa lay the joint mass of India, Antarctica and Australia.

Wegener gave numerous arguments for his theory. The most important ones were based on fossils. Fossils found in Antarctica and Patagonia made clear that the climate in these regions, now very cold, had once been tropical or subtropical. Moreover, the places where certain fossils were found in different continents turned out to be exactly contiguous when these continents were joint together according to their coastlines, as shown in the figure on the next page. The same was true for coal-beds and glacial deposits in different continents.

Wegener's theory was not received well, however. The continents were considered to be huge masses of land, resting on a lower layer of basalt rock called the crust. Therefore, to have the continents move all over the earth required a mechanism that could develop the force necessary to transport the continents over the crust. Wegener suggested that changes in the configuration of the sun and the moon and the coriolis forces due the earth's rotation could do the job, but this was not taken very seriously. The death blow to the theory was struck when it was shown that, as a result of their movement over the crust, the continents would crumble away, such that they should by now have disappeared altogether if they had been moving for the length of time supposed by Wegener's theory. Therefore it was concluded that Wegener's theory, holding that the positions of the continents change considerably and continuously on a geological time scale, was false.



The finding places of four fossils of organisms living during the Trias period (250-200 million years ago) are contiguous across up to five of the current continents if these are regrouped in a way that is suggested by their coastlines.

The theory was shelved until around 1960 systematic measurements started to be performed on the rock forming the bottom of the Atlantic Ocean, in particular on the direction of the magnetic field of the earth as it was 'frozen in' when the rock was formed. These measurements made clear that the bottom of the Atlantic Ocean is very young (on a geological time scale), originates in the middle and grows both in the direction of America and of Europe and Africa. This discovery formed the basis for the – currently universally accepted – theory of 'plate tectonics'. According to this theory, the continents move across the earth jointly with a segment of the crust and of the next layer, the upper mantle, directed by convection currents in the underlying, more or less fluid lower mantle of the earth. Rifts or faults divide the earth's crust and upper mantle into a number of plates that move relative to one another. Some of these plates increase in size because they grow at their edge, while others are gradually destroyed because their edge is forced downward and melts away. Scientists now accept continental drift as a fact.

This leads to the problem of understanding by what argument Wegener's theory was rejected in the 1930s, and to recognize what went wrong in this reasoning. First of all it is important to state Wegener's theory precisely in the form in which it is now accepted. I take it to be the theory that the continents move with respect to each other such that they cover distances of thousands of kilometers on a geological time scale of millions of years. To formalize the reasoning, let *w* be the theory of Wegener, and let *e* be the totality of the evidence, i.e. all the observations that Wegener on contiguous fossil founds, coal deposits, and so forth. For a start, let's emphasize that no deductive argument allows us to conclude the truth of *w* from the evidence. If the theory is true, we can predict *e* given a suitable empirical condition *c*, being that the continents have been moving for a sufficient time for fossil founds and coal deposits extending over various continents to have broken up. So we have 'If *w* is true, then if condition *c* obtains, we observe *e*', but even if *w*, *c* and *e* are so related, no deductively valid argument leads from *c* and *e* as premises to *w* as a conclusion. The falsification of *w*, on the other hand, seemed to be possible on the basis of a valid argument. Note, first of all, that one may be tempted to say that the theory was falsified by the fact that Wegener was unable to specify a plausible mechanism to propel the continents. Such an argument seems to fit the hypothetico-deductive method. If the theory is true, there must be a propelling mechanism *p*, leading to a first premise 'If *w* is true, then there is a propelling mechanism *p*'. If we add as a second premise 'There is no mechanism *p*', then it follows deductively that *w* is false. However, this means that the falsity of *w* is guaranteed by the falsity of 'There is a mechanism *p*'. But would we be justified in taking the premise 'There is no propelling mechanism *p*' to be true? Surely the fact that Wegener was not able to specify a plausible mechanism does not imply that there is no mechanism. The absence of a propelling mechanism is not an observation, like the absence of an electric current or the failure of a liquid to turn red.

The valid reasoning scheme of falsification guarantees the truth of the conclusion – hypothesis so-and-so is false – on the basis of the truth of the premises. Therefore the premises must be of the kind that we can be convinced of their truth: they must be empirical statements, capable of being verified directly by observation in nature or the laboratory. The failure to specify a propelling mechanism for the continents was not of this sort. But there was a statement, considered to follow from Wegener's theory, that was of the right sort. If the condition *c* obtains that the continents have in fact been moving during millions of years over thousands of kilometers, then the continents must have crumbled away. Indicate this prediction by *a*: the continents have crumbled away. Formally the prediction runs 'If *w* is true, then if *c* is the case, *a* is observed to be the case'. We accept that *c* is true (it must be for the theory to be compatible with the evidence collected by Wegener, so certainly Wegener and his followers must accept its truth) but we also know that *a* is false: the continents are all still there, and largely in the original size if, again, the theory is to be compatible with the other evidence. Thus the falsity of *w* follows deductively from the premises 'If *w* is true, then if *c* is the case, *a* is observed to be the case' and '*c* is the case but *a* is observed not to be the case'.

However, in deriving the prediction *a* from *w* and *c*, a crucial supposition was made. The continents are like huge masses of, mainly, basalt resting on top of the hard crust that covers the entire globe. Underneath this crust and the upper mantle, the material of the earth is in a highly viscous fluid state. The implicit assumption was that if the continents move, they move over the crust. The crumbling away then resulted from the friction between continent and crust. Introduce *m* for: the continents move over the crust. Then the correct hypothetico-deductive argument runs: if 'If *w* is true, then if *c* is the case, *a* is observed to be the case' and '*c* is the case but *a* is observed not to be the case', therefore: '*w* is false or *m* is false'. If we accept both the truth of *c* and the falsity of *a*, then either Wegener's theory is false or the assumption that the continents move over the crust is false, or possibly both. It is possible therefore to continue to accept *w*, if we recognize that *m* may well be false. And indeed, the theory of continental drift as it emerged from the measurements of the1960s contained the explicit denial of *m*: the earth's crust is not a single solid shell but is divided into a number of plates, and the continents do not move over the crust but move by riding piggy-back on the moving plates. In the course of this process, in their own way the plates indeed crumble away at some edges, whereas they grow at other ones.

5. *The poverty of deductive reasoning*

This completes the investigation to what extent reasoning in science can be reconstructed as applying deductively valid arguments. Deductive reasoning does not allow to say anything in favour of a theory's truth, irrespective of the number of occasions where the empirical facts were in agreement with the theory's predictions. Deductive reasoning does allow the firm conclusion of a theory's falsity, in case an observation is not in agreement with the theory's prediction. However, this applies without caveats only to the lowest-level empirical generalizations like 'all metals conduct electricity', which are stated in terms of directly observable quantities. The greater the level of generality of a theory, the larger the number of assumptions necessary to derive empirical statements from it. If these empirical statements conflict with the observations, then only the falsity of the totality of theory and assumptions can be concluded, but nothing specific follows with respect to the theory itself. This severely restricts the applicability of this form of argument.

It must be concluded, therefore, that deductive reasoning is not fit to model reasoning with respect to scientific knowledge. Scientists repeatedly take the corroboration of predictions as evidence for the truth of those theories. Much more than the mere number of corroborated predictions, the level of sophistication and unexpectedness of the prediction is considered important. When Niels Bohr formulated his theory of the atom, with its quantized electron orbits, in 1913 he was able to calculate that specific frequencies in the light emitted by hydrogen atoms and similar frequencies in the light emitted by ionized helium atoms should differ by a factor 4.00016, taking into account the mass ratios of hydrogen nuclei, helium nuclei and electrons. The actual ratio could be determined by measuring the absorption lines – Balmer lines for hydrogen and Pickering lines for helium – occurring in the spectra of stars, and was found to be equal to 4.000163. Learning of the agreement between the value calculated from Bohr's theory and the experimentally determined value, Einstein stated that Bohr's theory must be correct. We can safely assume that Einstein did not mean that the precise match between the prediction from Bohr's theory and the experimental results *proved* that the theory was true. The point is that scientists take such matches to provide very strong evidence for a theory, or a very good reason to accept a theory as correct. In fleshing this intuition out, deductive logic is shown to be of no use. The notion of the precision of a prediction, and of the relative unexpectedness of its agreeing exactly with the observed facts cannot even be represented by deductive logic.

Most scientists are also well aware of the other weakness of the hypothetico-deductive method, the weakness of falsification. The point is an illustration of the *Duhem-Quine thesis*, advocated in writing by the French physicist Pierre Duhem around 1900 and the American philosopher Willard Quine in the 1950s, but also by logical-empiricist philosophers like Rudolf Carnap and Otto Neurath in the 1930s. According to this thesis, a scientific theory cannot be tested in isolation; what is tested is always a conjunction of the theory and numerous auxiliary hypotheses and assumptions, or even, in the thesis's boldest formulation, what is tested is, again and again, the totality of our knowledge. Here as well, deductive logic can only indicate that somewhere in the conjunction of everything that is at issue something is wrong. It cannot be used to 'zoom in' and determine where exactly in this conjunction the blame can be put. Is the theory indeed false? Or should we rather retract some particular assumption that looks innocent but perhaps was never questioned before and will fail when it is? Of course each assumption could be treated as a hypothesis and a prediction could be derived from it that can be tested. However, this assumption can in its turn not be tested in isolation, and for many assumptions it may be very hard to derive predictions from them that can be tested in the current state of technology. Take the assumption crucial to a test of Galileo's law of free fall, namely that the air does not affect the rate of fall. An obvious prediction is that if there is no surrounding air, at least some falling bodies will fall differently, but this cannot be tested as long as no air pumps are available that can remove the air from containers that are large enough to allow a body to fall freely in it long enough, and as long as no containers are available that are large enough for this and in additional transparent enough to see what goes in inside. The first experiment showing that a metal ball and a feather fall with the same acceleration in a near vacuum was performed only in the second half of the 19th century, more than 200 years after Galileo published his law.

It seems we must accept, therefore, that a crucial aspect of scientific reasoning – deciding upon the epistemological status of generalizations, laws and theories – cannot be modelled by deductive logic. As a consequence, it seems we must accept that we will not be able to definitely establish the truth of scientific claims, since deductive logic *is* the inquiry of valid arguments, where the *truth* of the premises guarantees the *truth* of the conclusion. Nevertheless, we wish to uphold our conviction that scientists reason, because if there is no reasoning in science, where in life would there be any reasoning? Consequently, scientific reasoning has to be modelled, at least partly, as a form of non-deductive reasoning.

6. *Inductive vs. deductive reasoning*

Reasoning forms that are non-deductive but still intuitively acceptable are often jointly classified as forms of *inductive reasoning*, broadly conceived. Intuitively, an inductive argument is acceptable if the premises, provided they are true, *lend force to* the conclusion, or *support* the conclusion, without, however, guaranteeing its truth. The phrases 'lend force to' and 'support' are, of course, merely other, but equally vague, ways of expressing the intuition we have about the relation between premises and conclusion in an inductive argument. I have more to say on this below. The method of enumerative induction, by which the generalization that all metals conduct electricity is concluded from a great many observations of samples of metal that conduct electricity, is then an example of inductive reasoning in the broad sense.

How can we tell a deductive from an inductive argument? Suppose we have a valid deductive argument, and suppose the premises are all true. Therefore, necessarily, the conclusion is true. If we add premises to the ones already there, the conclusion must remain true, since the premises that guarantee the truth of the conclusion are all still there. In contrast, in an inductive argument the addition of premises may make it the case that the conclusion no longer 'follows', in the inductive sense, meaning that it is no longer supported by the premises or no longer receives force from the premises. The technical expression for this is that deductive reasoning is *monotonic* whereas inductive reasoning is *non-monotonic* or *defeasible*. Take the following example. Suppose we have *n* observations of samples of stuff that are made of metal and are silver-coloured. Suppose that the number of samples is fairly large, such that the argument

Object a_1 is made of metal and is silver-coloured
Object a_2 is made of metal and is silver-coloured
…
Object a_n is made of metal and is silver-coloured
___
All metal objects are silver-coloured

is deemed inductively acceptable. (Of course this argument is and remains deductively *in*valid.) But if the next object checked, a_{n+1}, is a piece of copper, then we add ‘Object a_{n+1} is made of metal and is not silver-coloured’ as a premise. The argument

Object a_1 is made of metal and is silver-coloured
Object a_2 is made of metal and is silver-coloured
…
Object a_n is made of metal and is silver-coloured
Object a_{n+1} is made of metal and is not silver-coloured

All metal objects are silver-coloured

is no longer inductively acceptable, irrespective of how exactly we flesh out the notion of inductively acceptable.

There is also an inductive analogue of prediction. Consider the argument:

Object a_1 is made of metal
Most metals are silver-coloured

[therefore,] Object a_1 is silver-coloured

We may find the conclusion that a_1 is silver-coloured acceptable, but it no longer is as soon as the premise ‘sample a_1 is made of copper’ is added, taking into account our knowledge that copper is not silver-coloured but reddish-yellow. The resulting set of three premises no longer supports the conclusion that a_1 is silver-coloured.

Summarizing, no inductive argument can ever count as a *proof* of its conclusion. Knowledge arrived at through inductive reasoning is always *fallible*. The importance of this insight can hardly be exaggerated.

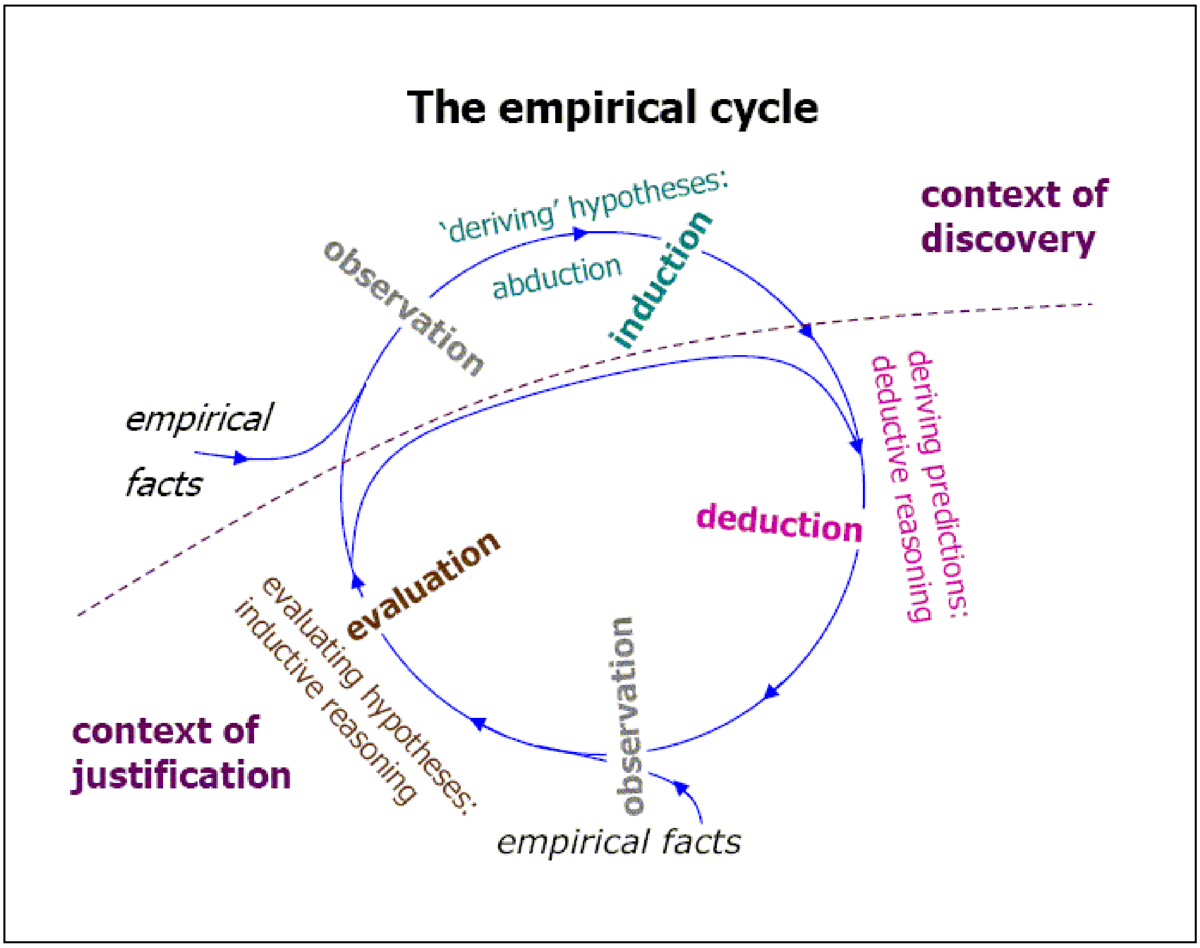
Colloquially, the fact that a conclusion can stop to ‘follow’ inductively from some premises, i.e. to be supported by them, as soon as the list of premises is extended reflects the fact that the conclusion of an inductive argument goes ‘beyond the premises’. It says something more than what is said by the conjunction of the premises. And what it says ‘in addition’ to the premises may be incompatible with what is said by an additional premise, or by what is implied by it in conjunction with our knowledge. The conclusion of a deductive argument, in contrast, does not go beyond the premises; what it says is already contained in the totality of the premises. This remains true, even if a premise is added that is incompatible with the existing premises. This is sometimes put as saying that deductive arguments are uninformative. This however, is misleading. A mathematical proof is a deductive argument, but mathematical proofs are seldom judged to be uninformative. It may come as quite a surprise that a particular statement is in fact contained in the conjunction of a particular set of statements.

7. *The empirical cycle; deduction, induction and abduction*

Although scientific reasoning cannot be just deductive, neither is it exclusively inductive. Let’s review the various reasoning forms that are involved in work in any empirical science. The various activities make up a process that is referred to as the empirical cycle. In the empirical cycle, five phases can be distinguished. Two of them are empirical activities. First,

at the beginning of it all, there is the taking in of unsolicited facts, things that come to be known by just looking around and observing the world or by ‘toying’ with experimental set-ups in a laboratory, out of curiosity, or things observed accidentally while doing experiments aimed at other matters. This first observational phase is followed by a reasoning phase: laws, theories, mechanisms are proposed that would account for the observed facts, unify them in some respect or explain them. Next, there comes another reasoning phase: once the laws, theories or mechanisms are articulated, consequences are derived from them, predictions concerning other facts that we must expect to obtain in particular circumstances if we are right about the theories and mechanisms. This phase is followed by another phase of activity: this time a meticulous search for the circumstances in nature in which the predicted fact should be observable, or the careful construction of an experiment aimed at creating these circumstances in a laboratory, terminating in a record of what was actually observed. This second empirical phase is followed by another reasoning phase, a phase of evaluation: given our observations, and the extent to which they match or fail to match the predictions derived from the proposed law or theory, what is the status of the latter? Can we go on to consider it as satisfactory accounts of nature’s ways, perhaps even as being true, or must we drop it?
After this evaluation phase, the cycle is closed. In its smallest loop new predictions are derived from the theories, which are again tested against the empirical facts. In a larger loop the theories and mechanisms again enter the generation phase and are modified or reformulated, sometimes beyond recognition, occasionally made to match new unsolicited facts as well, before novel predictions are derived from them.

The three reasoning phases of the empirical cycle are characterized by distinct reasoning



forms. The reasoning in the *prediction phase* is deductive: if the theories or mechanisms are true, and if the circumstances for which we make the prediction obtain, than we must expect the predicted phenomenon to occur. The discussion of sophisticated falsification has made clear that the derivation of predictions from theories always involves assumptions, but this serves to emphasize the deductive character of prediction; the assumptions are added to the premises in the derivation, or isolated within the theory at issue, in order to show that the truth of the prediction depends on the truth of these assumptions just as much as on the truth of the theory. Of course not all predictions are of the form that a particular effect will obtain with certainty, given the right circumstances. In the case of statistical predictions, what is predicted is not that some effect may occur, but that it will occur with some definite probability, or rather, that in a series of similar circumstances the effect will obtain in a definite fraction of the series.

The *evaluation phase*, on the other hand, cannot be deductive. That is what the above discussion of verification and falsification should have made clear. A realistic, or useful, framework for the evaluation of theories in the light of confirmed and failed prediction derived from them cannot take the form of assigning truth values to them. At least it cannot if we require it to be backed up by our framework for managing truth values, i.e. deductive logic. Instead, the reasoning in the evaluation process will have to be inductive in the broad sense, meaning that it does not assign truth values but something different. Prima facie, there seem two sensible possibilities. Inductive reasoning can go either by assigning a binary verdict similar to logic, classifying theories as ‘accepted’ and ‘rejected’, or in a continuous way, by assigning levels of support or lack of support, such that a theory can be assigned any level of support ranging from the maximum of support possible at one extreme of the spectrum to the minimum of support possible, or the maximum of lack of support, at the other extreme. Some proposals for setting up inductive reasoning in this way will be presented and discussed in detail in another chapter.

This leaves the *generation phase*. This phase is the least understood, and the reasoning pertaining to it discussed the least. Actually it is not clear that a discussion of this phase belongs to the problem of validating scientific claims, since validation is not the issue in this phase. Here the aim is to generate theories and mechanisms, in the first instance just as hypotheses, only to validate them once they are articulated in a precise enough form. This suggests that the empirical cycle falls apart in two quite distinct parts. In 1938 the philosopher Hans Reichenbach coined the names ‘context of discovery’ and ‘context of justification’ for these parts. The context of justification is the validation part of the cycle; it contains the prediction and evaluation phases. The reasoning going on in this part certainly is candidate for philosophical scrutiny, as it obviously matters to the cogency of our scientific knowledge what sort of conclusions are drawn from the confrontation of theories and empirical facts. Concerning the context of discovery, on the other hand, philosophy need perhaps say very little. For the logic of the evaluation phase, it is irrelevant whether a theory is the outcome of a particular reasoning process or has been ‘seen’ in a dream. The only thing that matters, the only thing that decides on its status as a scientific theory, is whether it predicts certain things to be the case and whether these predictions are borne out by the facts.

We can briefly put the general aims of the enterprise of science and technology as accepting as many claims as possible that are in fact true and as few claims as possible that are in fact false. Caring only about the logic of the context of justification may serve the latter of these two aims, but it will serve the former very poorly. We also know that very few of the theories and mechanisms currently accepted and applied in science and technology were received in a dream or vision. For promoting the growth of science and technology, an ability to understand and criticize the reasoning at work in theory formation and if possible to extend and improve it is highly desirable. The issue here is not to draw conclusions that amount to holding particular assertions true or sufficiently supported by the facts. The issue is rather to come up with assertions that are *candidates* for being held true or sufficiently supported by the facts. Whether they are so is to be determined in a later phase. These assertions need not emerge from the reasoning process with anything more than the status of being *prima facie* supported by whatever served as input of the process.

The reasoning underlying hypothesis formation (insofar as reasoning underlies it) is sometimes referred to as *abduction,* a term introduced by the American philosopher C.S. Peirce around 1900. This term suggests that abduction is a third form of reasoning, unlike both deduction and induction. Whether this is so, however, is highly questionable. Abduction may be better seen as a particular way of using the reasoning schemes of deductive logic and of inductive methods. In fact, abduction may turn out to be entirely analyzable in terms of deductive logic. For example, as mentioned above, the reasoning scheme of enumerative induction can be made deductively valid if a premise is added stating that the future is like the past. If we apply the scheme of enumerative induction extended in this way to arrive at the *truth* of the conclusion, the move will not help us, because, even though the truth of the premises (the statements reporting all or previous observations of As being Bs plus the statement that future observations will be similar) now guarantees the truth of the conclusion (the statement that all As are Bs) we have no access at all to the truth of the crucial premise that the future is like the past. For abduction, however, this is irrelevant, for we are only interested in acquiring the statement that all As are Bs *as a statement,* not as a true statement. Therefore, abduction is perhaps merely deductive logic combined with the strategic implementation of sweeping and unverifiable but highly productive premises. The fact that we have used premises claiming that the future is like the past or that the universe is simple so often with success may itself be seen as evidence that the universe is indeed uniform or simple, but that is another matter. A matter that one can consider lying beyond the concerns of science and technology, moreover, since it can be questioned whether there is any further use for these claims apart from there purely instrumental role in abductive reasoning.