Provisoes: A Problem concerning the Inferential Function of Scientific Theories
Author(s): Carl G. Hempel
Source: *Erkenntnis (1975-)*, Vol. 28, No. 2 (Mar., 1988), pp. 147-164
Published by: Springer
Stable URL: http://www.jstor.org/stable/20012141
Accessed: 30/01/2010 16:35

CARL G. HEMPEL

PROVISOES: A PROBLEM CONCERNING THE INFERENTIAL FUNCTION OF SCIENTIFIC THEORIES*[1]

1. INTRODUCTION

The principal goal and the proudest achievement of scientific inquiry is the construction of comprehensive theories which give us an understanding of large classes of empirical phenomena and enable us to predict, to retrodict, and to explain them.

These various functions of theories are usually regarded as having the character of inferences which lead, by way of theoretical principles, from sentences expressing initial and boundary conditions to statements describing the occurrences to be predicted, retrodicted, or explained.

In this paper, I propose to examine a basic difficulty which faces this inferential construal of scientific theorizing and which has implications for some central issues in the philosophy of science. I will first present the problem by reference to a purely deductivist conception of theoretical reasoning and will then broaden its scope.

2. THE STANDARD DEDUCTIVIST MODEL

The best-known precise elaboration of a deductivist conception is provided by the so-called standard empiricist construal of theories and their application. It views a theory T as characterizable by an ordered pair consisting of a set C containing the basic principles of the theory and a set I of interpretative statements:

(1) $T = \langle C, I \rangle$

The sentences, or formulas, of C serve to characterize the specific entities and processes posited by the theory (e.g., elementary particles and their interactions) and to state the basic laws to which they are assumed to conform. These sentences are formulated with the help of a theoretical vocabulary, V_C, whose terms refer to the kinds and characteristics of the theoretical entities and processes in question.

The sentences of the interpretative set I serve to link the theoretical

Erkenntnis **28** (1988) 147–164.

scenario represented by C to the empirical phenomena to which the theory is to be applied. These phenomena are taken to be formulated in a vocabulary V_A which is antecedently understood, i.e., which is available and understood independently of the theory. Thus, the sentences of I are said to provide partial interpretations, though not necessarily full definitions, of the theoretical terms in V_C by means of the antecedently understood terms of V_A. So-called operational definitions and reduction sentences in Carnap's sense may be viewed as special kinds of interpretative sentences.

By way of a simple example, assume that T is an elementary theory of magnetism whose theoretical vocabulary V_C contains such terms as 'magnet', 'north pole', 'south pole', and whose theoretical principles include the laws of magnetic attraction and repulsion and the law that the parts of a magnet are magnets again, while the class I includes some operational criteria for the terms of V_C.

Consider now the following application of the theory. From the sentence "b is a metal bar to which iron filings are clinging" (S^1_A), by means of a suitable operational criterion contained in the set I, infer "b is a magnet" (S^1_C); then, by way of theoretical principles in C, infer "If b is broken into two bars b_1 and b_2, then both are magnets and their poles will attract or repel each other" (S^2_C); finally, using further operational criteria from I, derive the sentence "If b is broken into two shorter bars and these are suspended by long thin threads close to each other at the same distance from the ground, they will orient themselves so as to fall into a straight line" (S^2_A). (Note that V_A is here taken to contain not only predicates like 'metal bar', but also individual constants such as 'b'.)

The basic structure thus attributed to a theoretical inference is suggested by the following schema, in which the notation $P \xrightarrow{Q} R$ is to indicate that R can be inferred from P by using sentences from Q as additional premisses.

$$(2)\qquad \begin{array}{ccc} S^1_C & \xrightarrow{\;C\;} & S^2_C \\ I\;\Big\uparrow & & \Big\downarrow\; I \\ S^1_A & & S^2_A \end{array}$$

Thus, if the inferential steps in question are indeed all deductive, then the theory provides a deductive inference bridge leading from one V_A-sentence, through the theoretical realm of C, to another V_A-sentence. More precisely: S_A^1 in combination with the theory T deductively implies S_A^2; this, in turn, is tantamount to saying that T deductively implies a corresponding V_A-sentence, namely, the conditional $S_A^1 \supset S_A^2$.

Carnap and other logical empiricists assumed that the vocabulary V_A, which serves to describe the phenomena to be explained by the theory, consists of terms that are "observational" at least in a broad sense, i.e., that they refer to features of the world whose presence or absence can be established by means of more or less direct observation. In recognition of the difficulties that face the notion of observability, I want to avoid any such assumption here. Indeed, I want to provide specifically for cases in which, as often happens, the vocabulary V_A was originally introduced in the context of an earlier theory. All that the standard construal needs to assume is that the phenomena for which the theory is to account are described by means of a vocabulary V_A that is "antecedently available" in the sense that it is well understood and is used with high intersubjective agreement by the scientists in the field. The interpretative sentences in I may then be viewed as interpreting the new terms introduced by the theory, i.e., those in V_C, by means of the antecedently understood terms in V_A.

This deductivist construal[2] faces two basic difficulties. I will call them the problem of theoretical ascent and the problem of provisoes. Let me spell them out in turn.

3. THEORETICAL OR INDUCTIVE ASCENT

The first inferential step in the schematic argument about the bar magnet presupposes that with the help of interpretative sentences belonging to part I of the theory of magnetism, S_C^1 is deducible from S_A^1. Actually, however, the theory of magnetism surely contains no general principle to the effect that when iron filings cling to a metal bar, then the bar is a magnet. The theory does not preclude the possibility, for example, that the bar is made of lead and is covered with an adhesive to which the filings stick, or that the filings are held in place by a magnet hidden under a wooden board supporting the lead bar. Thus, the theory does not warrant a deductive transition from S_A^1 to S_C^1. It is more plausible to assume that the theory contains an interpretative principle which is the converse of the one just considered, namely that if a bar is a magnet, then iron filings will cling to it. But even this is not strictly correct, as will be argued shortly.

Hence, the transition from S_A^1 to S_C^1 is not deductive even if the entire theory of magnetism is used as an additional premiss. Rather, the transition involves what I will call *inductive or theoretical ascent*, i.e., a transition from a data sentence expressed in V_A to a theoretical hypothesis S_C^1 which, by way of the theory of magnetism, would explain what the data sentences describes.

This illustrates one of the two problems mentioned before that face a strictly deductivist construal of the systematic connections which a theory establishes between V_A-sentences, i.e., between sentences describing empirical phenomena in terms of V_A. This problem has been widely discussed and various efforts have been made to resolve it by constructing theories of inductive reasoning that would govern such theoretical ascent. I will not consider those efforts here, but will rather turn to the problem of provisoes, which has not, it seems to me, been investigated in the same detail.

4. PROVISOES

Consider the third step in our example, the transition from S_C^2 to S_A^2. Again, the theory of magnetism does not provide interpretative hypotheses which would turn this into a strictly deductive inference. The theory clearly allows for the possibility that two bar magnets, suspended by fine threads close to each other at the same level, will not arrange themselves in a straight line; for example, if a strong magnetic field of suitable direction should be present in addition, then the bars would orient themselves so as to be parallel to each other; similarly, a strong air current would foil the prediction, and so forth.

The theory of magnetism does not guarantee the absence of such disturbing factors. Hence, the inference from S_C^2 to S_A^2 presupposes the additional assumption that the suspended pieces are subject to no disturbing influence or, to put it positively, that their rotational motions are subject only to the magnetic forces they exert upon each other.

Incidentally, the explanatory inference mentioned a moment ago, from S_C^1 to S_A^1, presupposes an analogous tacit premiss and thus is not deductive.

I will use the term '*provisoes*' to refer to *assumptions* of the kind just illustrated, *which are essential, but generally unstated, presuppositions of theoretical inferences*.

Provisoes are presupposed also in ostensibly deductive inferences that lead from one V_C-sentence to another. This holds, for example, in the inference from S_C^1 to S_C^2 in the case of the magnet: for if the breaking of the magnet takes place at a high temperature, the pieces may become demagnetized.

Or consider the application of the Newtonian theory of gravitation and of motion to a system of physical bodies like our solar system. In predicting, from a specification of the state of the system at a time t_0, subsequent changes of state, the basic idea is that the force acting on any one of the bodies is the vector sum of the gravitational forces exerted on it by the other bodies in accordance with the law of gravitation. That force then determines, via the second law of motion ($f = ma$), the resulting change of velocity and of position for the given body. But the quantity f in the second law is understood to be the *total* force acting on the given body; and the envisaged application of the theory therefore presupposes a proviso to the effect that the constituent bodies of the system are subject to no forces other than their mutual gravitational attraction. This proviso precludes not only gravitational forces that might be exerted by bodies outside the system, but also any electric, magnetic, frictional, or other forces to which the bodies in the system might be subject.

The absence of such forces is not, of course, vouchsafed by the principles of Newton's theory, and it is for this reason that the proviso is needed.

5. ESCAPE BY INTERPRETATIVE OF PROBABILISTIC FORM?

The foregoing considerations show in particular that when a theory contains interpretative sentences in the form of explicit definitions or of Carnapian reduction chains based on the antecedent vocabulary, the applicability of these sentences is usually subject to the fulfillment of provisoes; they cannot be regarded as unequivocal complete or partial criteria of applicability for theoretical expressions.

This thought might suggest a construal of the interpretative sentences of a theory as expressing only probabilistic rather than strictly general connections between theoretically described states or events and certain associated manifestations, or indicator phenomena, described in antecedently available terms. Such a construal might seem to come closer to scientific usage and at the same time to obviate the need for provisoes: for with probabilistic interpretation sentences, a theory would establish at best probabilistic connections between V_A-sentences. And what would otherwise appear as occasional violations of provisoes would be automatically anticipated by the merely probabilistic character of the theoretical inferences.

Interpretative sentences of probabilistic form have in fact been envisaged by several writers. Carnap did so already in his (1956) essay 'The Methodological Character of Theoretical Concepts', which is, I think, his earliest full presentation of the standard empiricist construal of theories. He argues there that many terms functioning in scientific theories cannot be regarded as linked to antecedent terms ("observational terms") by interpretative sentences ("rules of correspondence") of strictly universal form. For such sentences would specify strictly necessary or sufficient observational conditions of applicability for the theoretical terms, whereas scientists, Carnap argues, will treat such conditions not as strictly binding, but as qualified by an "escape clause" to the effect that the observational criteria hold "unless there are disturbing factors", or "provided the environment is in a normal state".[3] Such escape clauses clearly have the character of provisoes in the sense adumbrated earlier. Carnap views them as probabilistic qualifiers functioning in interpretative sentences for theoretical terms. These sentences would state probabilistic rather than strictly necessary or sufficient connections between theoretical expressions and V_A-sentences. Indeed, while Carnap countenances dispositional terms, linked to V_A by strict reduction chains, he suggests that the terms characteristic of scientific theories have only probabilistic links to the observational basis.[4]

But while Carnap thus explicitly eschews a purely deductivist construal of the inferential function of theories, he does not specify the form of the probabilistic interpretation sentences he envisages. Indeed, in response to a proposal by Pap[5] concerning probabilistic reduction sentences, Carnap remarks: "it seems to me that for the time being the problem of the best form for [interpretative sentences] has not yet been sufficiently clarified".[6]

However that may be, a probabilistic construal of provisoes faces the difficulty that scientific theories do not, in general, provide probabilistic laws that would obviate the need for provisoes.

Consider, for example, the interpretative sentences that would be required for the term 'magnet'. They would have to take the form "In cases where iron filings stick to a metal bar, the probability of the bar being a magnet is p_1"; or, for inferences in the opposite direction: "Given that a metal bar is magnetic, the probability that iron filings will cling to it is p_2". But surely, the theory of magnetism contains no sentences of this kind; it is a matter quite beyond its scope to state how frequently air currents, disturbing further magnetic fields, or other factors will interfere with the effect in question. It seems to me that no scientific theory provides probabilistic interpretation statements of this sort, whose application is not itself subject to provisoes.

The same basic consideration applies also, I think, where no well-developed and sharply formulated theories are available; for example, probabilification cannot avoid the need for provisoes in the application of theoretical sentences linking psychological states or events to their behavioral manifestations.

6. SOME CONSEQUENCES OF THE NEED FOR PROVISOES

The conclusion that a scientific theory even of non-probabilistic form does not, in general, establish deductive bridges between V_A-sentences has significant consequences for other issues in the philosophy of science.

I will briefly indicate four of these: (a) the idea of falsifiability, (b) the significance of so-called elimination programs for theoretical terms, (c) the instrumentalist construal of scientific theories, and (d) the idea of "the empirical content" of a theory.

(a) *Falsifiability*

One obvious consequence of the need for provisoes is that normally a theory is not falsifiable by V_A-sentences; otherwise, it would deductively imply the negations of the falsifying V_A-sentences, which again are V_A-sentences.

This consideration differs from the Duhem-Quine argument that individual hypotheses cannot be falsified by experiential findings because the deduction from the hypothesis of falsifying V_A-sentences requires an extensive system of background hypotheses as additional premisses, so that typically only a comprehensive set of hypotheses will entail or contradict V_A-sentences. The argument from provisoes leads rather to the stronger conclusion that even a comprehensive system of hypotheses or theoretical principles will not entail any V_A-sentences because the requisite deduction is subject to provisoes.

Note that a proviso as here understood is not a clause that can be attached to a theory as a whole and vouchsafe its deductive potency by asserting that in all particular situations to which the theory is applied, disturbing factors are absent. Rather, a proviso has to be conceived as a clause which pertains to some particular application of a given theory and which asserts that in the case at hand, no effective factors are present other than those explicitly taken into account.

(b) *Elimination programs for theoretical terms*

The need for provisoes also has a bearing on the so-called elimination programs for theoretical terms. These programs are of particular significance for philosophical qualms about the use, in scientific theories, of terms that are not explicitly defined by means of an antecedently understood vocabulary.

The ingenious and logically impeccable methods designed by Ramsey and by Craig[7] circumvent these qualms by showing that the use of theoretical expressions can always be avoided in the following sense: If a theory T consisting of two sentence classes C and I as characterized earlier does yield deductive connections between certain V_A-sentences, then it is possible to formulate a corresponding theory (class of sentences) T_A such that

(i) T_A is expressed in terms of V_A alone
(ii) T_A is logically implied by T
(iii) T_A entails '$S_A^1 \supset S_A^2$' (and in this sense establishes a deductive bridge from S_A^1 to S_A^2) if and only if T entails '$S_A^1 \supset S_A^2$'.[8]

If the function of a theory is taken to consist in establishing deductive bridges among V_A-sentences, then the theory T_A, which avoids the use of theoretical terms, might be called functionally equivalent to the theory T. This result might suggest the reassuring conclusion that, in principle, the use of theoretical expressions can always be avoided without any change in the "empirical content" of a theory as it is expressed by the class of V_A-sentences deducible from it, and that talk in terms of theoretical expressions is just a convenient *façon de parler* about matters that are fully expressible in the antecedently understood vocabulary V_A. Analogously, it may seem that all the problems about theoretical ascent and provisoes simply disappear if T is replaced by its functional equivalent T_A.

This impression is illusory, however. For a theory T_A constructed from T in the manner of Ramsey or of Craig yields deductive connections between V_A-sentences if and only if T yields such connections: and scientific theories do not, in general, satisfy this condition. The need for provisoes precludes the general avoidability of theoretical expressions by those elimination methods.

The verdict does not hold, however, if the provisoes qualifying the inferential applications of a theory are themselves expressible in the antecedent vocabulary. For if P_A is such a proviso governing the transition, by means of T, from S^1_A to S^2_A, then T entails the sentence $(P_A \cdot S^1_A) \supset S^2_A$ and thus establishes a deductive bridge between two V_A-sentences.

But it seems that, in general, the requisite provisoes cannot be expressed in terms of V_A alone. In the case of the theory of magnetism referred to earlier, the provisoes may assert, for example, the absence of other magnetic fields, or of disturbing forces, etc., and will then require at least the use of terms from V_C in their formulation.

(*c*) *Provisoes and the instrumentalist perspective*

The preceding considerations analogously cast some doubt on the instrumentalist conception of theories as purely inferential devices which, from an input in the form of V_A-sentences, generate an output of other V_A-sentences. For the need for provisoes shows that theories do not render this service. In each particular case, the applicability of the theoretical instrument would be subject to the condition that the pertinent provisoes are fulfilled; and the assertion that they are fulfilled could not just be added to the input into the theoretical calculating machine, for that assertion would not generally be expressible in V_A.

Thus, if a theory is to be thought of as a calculating instrument that generates new V_A-sentences from given ones, then it must be conceived as supplemented by an instruction manual specifying that the instrument should be used only in cases in which certain provisoes are satisfied. But the formulation of those provisoes will make use of V_C and perhaps even of terms not contained in V_C. Thus, one has to check whether certain empirical conditions not expressible in V_A are satisfied: and that surely provides a tug away from instrumentalism and in the direction of realism concerning theoretical entities.

(d) Provisoes and "the empirical content" of a theory

Similar questions arise in regard to the notion of the experiential "cash value" or "empirical content" of a theory as represented by the set of all V_A-sentences entailed by the theory.

Note first, and incidentally, that thus construed the empirical content of a theory is relative to the vocabulary V_A that counts as antecedently available, so that one would properly have to speak, not of "the" empirical content of T, but of the V_A-content of T.

But the point here to be made is rather that usually a theory does not entail V_A-sentences and the proposed construal of empirical content misfires.

To be sure, there are some deductive theoretical inferences that presuppose no provisoes; for example, the inference, mediated by the law of gravitation, from a sentence S^1 specifying the masses and the distance of two bodies to a sentence S^2 specifying the gravitational attraction that the bodies exert upon each other.

But the further theoretical inference from S^2 to a sentence S^3 specifying the accelerations the bodies will undergo requires a proviso to the effect that no other forces act upon the bodies. If S^2 and S^3 are represented as theoretical sentences, then we have here an example of the need for provisoes not only in establishing theoretical inference bridges between V_A-sentences and V_C-sentences, but also in building such bridges between sentences expressed solely in terms of V_C. We will shortly return to this point.

7. FURTHER THOUGHTS ON THE CHARACTER OF PROVISOES

How might the notion of proviso be further illuminated? To say that provisoes are just *ceteris paribus* clauses is unhelpful, for the idea of a *ceteris paribus* clause is itself vague and elusive. "Other things being equal, such-and-so is the case". What other things, and equal to what? How is the clause to function in theoretical reasoning?

Provisoes might rather be viewed as *assumptions of completeness*. The proviso required for a theoretical inference from one sentence, S^1, to another, S^2, asserts, broadly speaking, that in a given case (e.g., in that of the metal bar considered earlier) no factors other than those specified in S^1 are present which could affect the event described by S^2.

For example, in the application of Newtonian theory to a double star it is presupposed that the components of the system are subject to no forces other than their mutual gravitational attraction and hence, that the specification given in S^1 of the initial and boundary conditions which determine that gravitational attraction is a complete or exhaustive specification of all the forces affecting the components of the system.

Such completeness is of a special kind. It differs sharply, for example, from that invoked in the requirement of complete or total evidence. This is an epistemological condition to the effect that in a probabilistic inference concerning, say, a future occurrence, the total body of evidence available at the time must be chosen as the evidential basis for the inference.[9]

A proviso, on the other hand, calls not for epistemic, but for ontic completeness: the specifics expressed by S^1 must include, not all the information available at the time (information which may well include false items), but rather all the factors present in the given case which in fact affect the outcome to be predicted by the theoretical inference; the factors in question might be said to be those which are "nomically relevant" to the outcome, i.e., those on which the outcome depends in virtue of nomic connections.

Consider once again the use of Newtonian theory to deduce, from a specification S^1 of the state of a binary star system at time t_1, a specification S^2 of its state at t_2. Let us suppose, for simplicity, that S^1 and S^2 are couched in the language of the theory; this enables us to leave on one side the problem of the inductive ascent from astronomical observation data to a theoretical re-description in terms of masses, positions, and velocities of the two objects.

The theoretical inference might then be schematized thus:

(3) $(P \cdot S^1 \cdot T) \rightarrow S^2$

where P is a proviso to the effect that apart from the circumstances specified in S^1 the two bodies are, between t_1 and t_2, subject to no influences from within or from outside the system that would affect their motions. The proviso must thus imply the absence, in the case at hand, of electric, magnetic, and frictional forces, of radiation pressure and of any telekinetic, angelic, or diabolic influences.

One may well wonder whether this proviso can at all be expressed in the language of celestial mechanics, or even in the combined languages of mechanics and other physical theories. At any rate, neither singly nor jointly do those theories assert that forces of the kinds they deal with are the only kinds by which the motion of a physical body can be affected. A scientific theory propounds an account of certain kinds of empirical phenomena, but it does not pronounce on what other kinds there are. The theory of gravitation neither asserts nor denies the existence of non-gravitational forces, and it offers no means of characterizing or distinguishing them.

It might seem, therefore, that the formulation of the proviso transcends the conceptual resources of the theory whose deductive applicability it is to secure. That, however, is not the case in the example at hand. For in Newton's second law, $f = ma$, 'f' stands for the *total* force impressed on the body; and our proviso can therefore be expressed by asserting that the total force acting on each of the two bodies equals the gravitational force exerted upon it by the other body; and the latter force is determined by the law of gravitation.[10]

But the application of the theory to particular cases is clearly subject again to provisoes to the effect that in computing the total force, all relevant influences affecting the bodies concerned have been taken into account.

When the application of a theory to empirical subject matter is schematically represented in the form (3) with the provisoes P as one of the premisses, it must be borne in mind that the language and the specific form in which P is expressed are left quite vague. The notation is not meant to be a sharp explication, but rather a convenient way of referring to the subject at issue in the context of an attempt to shed some further light on it.

Note that the proviso P does not include clauses to the affect that the establishment of S^1 has not been affected by errors of observation or measurement, by deceit or the like: that is already implied by the premiss S^1 itself, which trivially asserts that S^1 is true. The proviso is to the effect, not that S^1 is true, but that it states the *whole* truth about the relevant circumstances present.

Note further that the perplexities of the reliance on provisoes cannot be avoided by adopting a structuralist, or non-statement, conception of theories broadly in the manner of Sneed and Stegmüller.[11] That conception construes theories not as classes of statements, but as deductively organized systems of statement functions, which make no assertions and have no truth values. But such systems are presented as having empirical models; for example, the solar system might be claimed to be a model of a structuralist formalization of Newtonian celestial mechanics. But a formulation of this claim, and its inferential application to particular astronomical occurrences, again clearly assumes the fulfillment of pertinent provisoes.

8. METHODOLOGICAL ASPECTS OF PROVISOES

The elusive character of proviso clauses raises the question of how a theoretical inference of type (3) can be applied to particular occurrences, and more specifically, on what grounds proviso P may be taken to be satisfied or violated in specific cases.

There are circumstances that provide such grounds. If the theory T has strong previous support, but its application to a new case yields incorrect predictions S^2, then doubts may arise about S^1; but in the absence of specific grounds for such doubts, a violation of P – i.e., the presence of disturbing factors – may suggest itself. If this conjecture can be expressed in the language of the theory T and replacing S^1 by a correspondingly modified sentence $S^{1'}$ yields successful predictions, then this success will constitute grounds for attributing the predictive failure of the original theoretical inference to a violation of its proviso clause.

Thus, the failure of Newton's otherwise highly successful theory to predict certain perturbations in the orbit of Uranus in terms of the gravitational attraction exerted on it by the sun and by the planets known before 1846 led to the conjecture of a proviso-violation, namely the assumption that Uranus was subject to the additional attraction of a hitherto unknown planet – a conjecture borne out by the subsequent discovery of Neptune.

Sometimes predictive failure of a theory is attributed to proviso-violations even though the presumably disturbing factors cannot be adequately specified.

Consider, for example, the controversy between Robert A. Millikan and Felix Ehrenhaft over the results of the famous experiments in which Millikan measured the rates at which small electrically charged oil drops rose and sank in the electric field between two horizontal electrically charged metal plates. From those rates he computed, by means of accepted theoretical principles, the size of the charges of those oil drops and found that all of them were integral multiples of a certain minimum charge *e*, whose numerical value he specified. Millikan presented his findings as evidence for the claim that electricity had an atomistic structure and that the atoms of electricity all had the specified charge *e*.

Ehrenhaft objected that in similar experiments, he had found individual charges which were not integral multiples of Millikan's value *e* and which, in fact, were often considerably smaller than *e*, suggesting the existence of "sub-electrons".[12] Ehrenhaft accordingly rejected Millikan's theoretical claims T on grounds of predictive failure.

Millikan replied in careful detail. Referring to difficulties he had encountered in his own work, he argued that Ehrenhaft's deviant results could be due to disturbing factors of various kinds. Among them, he mentions the possibility that tiny dust particles might have settled on the falling oil droplets, thus changing the total force acting on them; the possibility that evaporation might have reduced the mass of an observed drop; the possibility that the strength of the electric field might have decreased as a result of battery fatigue, and so forth.

Ehrenhaft repeated his experiments, taking great pains to screen out such disturbing factors, but he continued to obtain deviant findings. The sources of these deviations have never been fully determined; in fact, Ehrenhaft's results turned out not to be generally reproducible. Millikan's ideas, on the other hand, were sustained in various quite different applications. Thus, eventually Ehrenhaft's claims were gradually disregarded by investigators in the field, and Millikan won the day and the Nobel prize.

Interestingly, as has been pointed out by Holton,[13] Millikan himself had recorded in his laboratory diaries several sets of quite deviant measurements, but he had not published them, attributing them to disturbing factors of various kinds and sometimes not even offering a guess as to the source of the deviation.

But evidently, it cannot be made a *general* policy of scientific research to attribute predictive failures of theoretical inferences to the violation of some unspecified proviso; for this "conventionalist stra-tegem", as Popper has called it, would deprive a theory of any predictive or explanatory force.[14]

I think that at least in periods of what Kuhn calls normal science a search for disturbing influences will consider only factors of such kinds as are countenanced by one or another of the currently accepted scientific theories as being nomically relevant to the phenomena under consideration.

Thus, if a prediction based on Newtonian mechanics fails, one might look for disturbing gravitational, electric, magnetic, and frictional forces and for still some other kinds, but not for telekinetic or diabolic ones. Indeed, since there are no currently accepted theories for such forces, we would be unable to tell under what conditions and in what manner they act; consequently, there is no way of checking on their presence or absence in any particular case.

The mode of procedure just mentioned is clearly followed also in experiments that require screening-out of disturbing outside influences – for example, in experimental studies of the frequency with which a certain kind of subatomic event occurs under specified conditions. What outside influences – such as cosmic rays – would affect the frequency in question, and what shielding devices can serve to block them and thus to ensure satisfaction of the relevant proviso, is usually determined in the light of available scientific knowledge, which again would indicate no way of screening out, say, telekinetic influences.

If a theory fails to yield correct predictions for a repeatable phenomenon by reference to factors it qualifies as relevant, then certain changes within that theory may be tried, introducing a new kind of nomically relevant factor. Roentgen's discovery of a photo-graphic plate that had been blackened while lying in a closed desk drawer is, I think, a case in point; it led to the acknowledgment of a new kind of radiation.

Finally, persistent serious failures of a theory may lead to a rev-olution in Kuhn's sense, which places the phenomena into a novel theoretical framework rather than modify the old one by piecemeal changes. In this case, the failures of the earlier theory are not attributed to proviso violations; indeed it is quite unclear what such an attribution would amount to.

Consider a theoretical inference that might have been offered some 250 years ago on the basis of the caloric fluid theory of heat or the phlogiston theory of combustion. The relevant provisoes would then have to assert, for example, that apart from the factors explicitly taken into account in the inference, no other factors are present that affect, say, the flow of caloric fluid between bodies or the degree of dephlogistication of a body. But from our present vantage point, we have to say that there are no such substances as caloric fluid or phlogiston, and that therefore there could be no proper proviso claim of the requisite sort at all.

And yet, it appears that the claims and the inferential applications of any theory have to be understood as subject to those elusive provisoes.

There is a distinct affinity, I think, between the perplexing questions concerning the appraisal of provisoes in the application of scientific theories and the recently much discussed problems of theory choice in science.

As Kuhn in particular has argued in detail, the choice between competing theories is influenced by considerations concerning the strength and the relative importance of various desirable features exhibited by the rival theories; but these considerations resist adequate expression in the form of precise explicit criteria. The choice between theories in the light of those considerations, which are broadly shared within the scientific community, is not subject to, nor learned by means of, unambiguous rules. Scientists acquire the ability to make such choices in the course of their professional training and careers, somewhat in the manner in which we acquire the use of our language largely without benefit of explicit rules, by interaction with competent speakers.

Just as, in the context of theory choice, the relevant idea of superiority of one theory to another has no precise explication and yet its use is strongly affected by considerations shared by scientific investigators, so in the inferential application of theories to empirical contexts, the idea of the relevant provisoes has no precise explication, yet it is by no means arbitrary and its use appears to be significantly affected by considerations akin to those affecting theory choice.

NOTES

* This article has grown out of a paper read in November 1980 at a workshop held under the auspices of the Center for Philosophy of Science at the University of

Pittsburgh. The present, much revised, version was written for inclusion in a volume, to be published by the University of California Press, which is to contain the proceedings of that workshop.

Pending the completion of that project, which has been considerably delayed, the article appears here with the consent of the editors of the proceedings.

[1] This paper is based upon work supported by National Science Foundation Grant No. SES 80-25399.

[2] I have limited myself here to a schematic account of those features of the empiricist model which are of relevance of the problems subsequently to be discussed. For fuller expositions and critical discussions, and for references to the extensive literature, see, for example, Carnap (1956, 1966, chaps. 23–26); Feigl (1970); Hempel (1958, 1969, 1970); Putnam (1962); and Suppe (1974), a comprehensive study which includes a large bibliography.

[3] Carnap (1956), p. 69.

[4] Cf. Carnap (1956), pp. 49, 72.

[5] Pap (1963), section II.

[6] Carnap (1963), p. 950.

[7] For details, see Ramsey (1931, 'Theories', section IXA); Carnap (1966, 'The Ramsey Sentence', chap. 26); Craig (1956); Putnam (1965); Hempel (1965, pp. 210–17).

[8] The theory T_A obtainable by Ramsey's method is quite different, in other respects, from that generated by Craig's procedure. But the differences are irrelevant to the point here under discussion.

[9] Cf. Carnap (1950), pp. 211–13; 494.

[10] I am indebted to Michael Friedman for having pointed this out to me.

[11] Cf. Sneed (1979), Stegmüller (1976), especially chap. 7.

[12] Millikan gives a detailed account of his investigations in Millikan (1917); Ehrenhaft's claims are discussed in Chapter VIII. The controversy is examined in a broader scientific and historical perspective in Holton (1978).

[13] Holton (1978, esp. pp. 58–83).

[14] See, for example, Popper (1962, pp. 33–39); Stegmüller (1976, chap. 14).

REFERENCES

Carnap, R.: 1950, *Logical Foundations of Probability*, University of Chicago Press, Chicago.

Carnap, R.: 1956, 'The Methodological Character of Theoretical Concepts', in *Minnesota Studies in the Philosophy of Science*, vol. 1, H. Feigl and M. Scriven (eds.), University of Minnesota Press, Minneapolis, pp. 38–76.

Carnap, R.: 1963, 'Arthur Pap on Dispositions', in *The Philosophy of Rudolf Carnap*, P. A. Schilpp (ed.), Open Court, LaSalle, Illinois, pp. 947–52.

Carnap, R.: 1966, *Philosophical Foundations of Physics*, Basic Books, New York.

Craig, W.: 1956, 'Replacement of Auxiliary Expressions', *Philosophical Review* **65**, 38–55.

Feigl, H.: 1970, 'The 'Orthodox' View of Theories: Remarks in Defense as Well as Critique', in *Minnesota Studies in the Philosophy of Science*, vol. 4, M. Radner and S. Winokur (eds.), University of Minnesota Press, Minneapolis, pp. 3–16.

Hempel, C. G.: 1958, 'The Theoretician's Dilemma', in *Minnesota Studies in the Philosophy of Science*, vol. 2, H. Feigl, M. Scriven, and G. Maxwell (eds.), University of Minnesota Press, Minneapolis, pp. 37–98, reprinted in C. G. Hempel: 1965, *Aspects of Scientific Explanation*, The Free Press, New York, pp. 173–226.

Hempel, C. G.: 1969, 'On the Structure of Scientific Theories' in *Isenberg Memorial Lecture Series*, 1965–66, Michigan State University Press, East Lansing, pp. 11–38.

Hempel, C. G.: 1970, 'On the 'Standard Conception' of Scientific Theories', in *Minnesota Studies in the Philosophy of Science*, vol. 4, M. Radner and S. Winokur (eds.), University of Minnesota Press, Minneapolis, pp. 142–63.

Holton, Gerald: 1978, 'Subelectrons, Presuppositions, and the Millikan-Ehrenhaft Dispute', in G. Holton, *The Scientific Imagination*, Cambridge University Press, Cambridge, pp. 25–83.

Millikan, Robert A.: 1917, *The Electron*, University of Chicago Press, Chicago, facsimile edition, 1963.

Pap, A.: 1963, 'Reduction Sentences and Disposition Concepts', in *The Philosophy of Rudolf Carnap*, P. A. Schilpp (ed.), Open Court, LaSalle, Illinois, pp. 559–97.

Putnam, H.: 1962, 'What Theories are Not', in *Logic, Methodology and Philosophy of Science*, E. Nagel, P. Suppes, and A. Tarski (eds.), Stanford University Press, pp. 240–51.

Popper, K.: 1962, *Conjectures and Refutations*, Basic Books, New York.

Putnam, H.: 1965, 'Craig's Theorem', *The Journal of Philosophy* **62**, 251–60.

Ramsey, F. P.: 1931, *The Foundations of Mathematics and Other Logical Essays*, edited by R. B. Braithwaite, Routledge and Kegan Paul, London; paperback reprint, 1960, Littlefield, Adams; Patterson, New Jersey.

Sneed, J. D.: 1979, *The Logical Structure of Mathematical Physics*, 2nd ed., D. Reidel, Dordrecht.

Stegmüller, W.: 1976, *The Structure and Dynamics of Theories*, Springer Verlag, New York.

Suppe, F.: 1974, 'The Search for Philosophic Understanding of Scientific Theories', in *The Structure of Scientific Theories*, F. Suppe (ed.), University of Illinois Press, pp. 3–232, (second edition, with an added "Afterword", 1977).

Manuscript received 1 June 1987

Department of Philosophy
Princeton University
Princeton, NJ08544
U.S.A.