Logic and the Philosophy of Science

Bas C. van Fraassen
Department of Philosophy
San Francisco State University

Abstract

While logic has sometimes tended to lead to oversimplification and abstraction, it has also made it possible to refine philosophical problems pertaining to science so as to give them rigor and precision, and in some cases, to solve them definitively. There are too many different cases to provide a helpful overview, so I will discuss several examples that I have found especially telling concerning the value of logic. I will take up two issues concerning definability and one issue in epistemology. They concern the problem of understanding theoretical terms in physics and what is known as the problem of old evidence.

1   The Problem of 'Implicit' Definability

How theoretical terms are related to what we can observe and measure has been a recurrent problem in philosophy of science. When the theoretical scene changes, new terms appear and to understand what they mean seems to require learning the new theories, as if they can only be understood 'from within' those new theories.

I will begin with a famous historical example. In the 17th century, Cartesians considered Newton's introduction of the new concepts of mass and force a return to the 'occult qualities' of the medievals. In the 19th century, however, there were sustained efforts to provide reductive accounts of those concepts. Mach's work is the best known. His attempt and its difficulties presaged the wider ranging controversies about theoretical terms in our own time, which we will address in turn.

1.1   Attempted definitions of mass

In the context of classical physics, all measurements are reducible to series of measurements of time and position, so we may designate as basic

1. The Problem of 'Implicit' Definability

1.1. Attempted definitions of mass
   a. Critique of the definition
   b. Bressan: Necessity rather than counterfactuals
   c. Alternative approaches to mass in mechanics
1.2 Attempted eliminations of theoretical terms
   a. Hilbert's introduction of 'implicit definition'
   b. Quine's 'Implicit definition sustained'
   c. Keeping something fixed: Winnie's rejection
   d. Lewis on the definition of theoretical terms
   e. The relevance of Beth's theorem

2. The Problem of Old Evidence
   2.1 Evidence, confirmation, and probability
   2.2 Responses to the problem of old evidence
   2.3 Postulating new 'logical' evidence
   2.4 Probabilist versions of implication and modus ponens
   2.5 The Conditional Proof Requirement

3. Conclusion

observables all quantities that are functions of time and position alone. Called the kinematic quantities, these include velocity and acceleration, relative distances and angles of separation. They do not include mass, force, momentum, kinetic energy (the dynamic quantities). To some extent the values of the latter can be calculated (within the theory) from the basic observables. That is precisely what inspired the many proposed 'definitions' of force and mass in the nineteenth century, and the more recent axiomatic theories of mechanics in which mass is not a primitive quantity.

The seminal text is Mach's proposed 'definition' of this concept within the theory, in a chapter called "Criticism Of The Principle Of Reaction And Of The Concept Of Mass" (Mach 1883: 264ff). He writes, in somewhat tentative fashion:

> If, however, mechanical experiences clearly and indubitably point to the existence in bodies of a special and distinct property determinative of accelerations, nothing stands in the way of our arbitrarily establishing the following definition:
>
> *All those bodies are bodies of equal mass, which, mutually acting on each other, produce in each other equal and opposite accelerations.*
>
> [...] That these accelerations always have opposite signs, that there are therefore, by our definition, only positive masses, is a point that experience teaches, and experience alone can teach. In our concept of mass no theory is involved; "quantity of matter" is wholly unnecessary in it; all it contains is the exact establishment, designation, and denomination of a fact. (pp. 266-7)

While the 'definition' is formulated in terms that are purely kinematic, Mach clearly realized that there is an empirical fact behind it, so to speak, in that at least the law of equality of action and reaction must be presupposed. But Mach's idea of a definition does not meet the standards developed since then.

1.1.1   Critique of the definition

As Patrick Suppes emphasized, if we postulate with Newton that every body has a mass, then mass is not definable in terms of the basic observables, not even if we take force for granted (cf. Suppes 1957: 298). For, consider, as simplest example, a model of mechanics in which a given body

has constant velocity throughout its existence. We deduce, within the theory, that the total force on it equals zero throughout. But every value for its mass is compatible with this information.

Could Mach possibly have missed this obvious point? It seems unlikely. It appears rather that his purpose was to present what in his book on the theory of heat was called a *coordination* (Mach 1896: 52). The concept of mass is introduced within the theory by specifying precisely what will count as a measurement of mass, by procedures that presuppose the empirical correctness of Newton's third law of action and reaction. These procedures are explored in some detail in a previous section (Mach 1883: 247ff.). What Mach thinks of as the axiomatizations of the empirical core and the general theory of classical mechanics is not the same enterprise as what 20th century logicians consider to be definition and axiomatization. But Mach's work has the virtue of focusing on the relation between the mathematics and experimental practice.

Nevertheless, Suppes has a point, and it is clear that if Mach wants to have the theory imply that every body has a mass, then he is involved in a modal or counterfactual assertion about *what would happen* under suitable, possible but not always actual, conditions.

What are those conditions, and to what extent do they determine the masses of bodies, relative to the theory? Here the seminal work was by Pendse (1937, 1939, 1940) to determine how much information about a body (possibly concerning a number of distinct times) would allow one to calculate its mass (cf. Jammer1964: 92-95). It appeared that *in almost all* cases, the kinematic data would determine the mass. In response to Suppes, Herbert Simon (1954, 1959, 1966) discussed a measure on the class of models of Newtonian particle mechanics, and proved that by that measure mass is definable *almost everywhere*. However, although this measure was presented as 'natural', we must acknowledge the infinity of exceptions this "almost everywhere" allows.[1] Such results are scant comfort for someone who wishes to eliminate mass as a primitive concept.

To sum up: there are models of mechanics (that is, worlds allowed as possible by this theory) in which a complete specification of the basic observable quantities does not suffice to determine the values of all the other quantities. Thus the same observable phenomena equally fit more than one distinct model of the theory.

1.1.2 Bressan: Necessity rather than counterfactuals

Use of counterfactual language can raise eyebrows even among the friends of modality. Is it possible at least in the present case to replace the counterfactuals with the alethic modalities, necessity, possibility, ...? To do so was the aim of Aldo Bressan (1973), who devised a subtle and rich account of the modalities and constructed formal proofs of adequacy for his axiomatization of mechanics. While I do not wish to discuss this work in detail, there is a problem concerning the specific formulation adopted by Bressan, and it concerns a quite general difficulty for the understanding of modalities in nature. In his formulation, Bressan asserts for each body U a conditional :

(If a certain experiment is performed on U then the outcome is real number ρ)

which he ranks as necessary, and as satisfied by a unique number ρ. Upon analysis, it then appears that this means that in all physically possible cases, this experiment upon U yields ρ.

What we need to ask then, however, is: what are the physically possible cases? They cannot be those logically possible cases that are compatible with the laws of mechanics, for the laws, being general, will not entail information about characteristics of individual, specific bodies. If to those laws we add factual information about U, phrased in purely kinematic terms, there will still be in general many alternatives left open. That is just our initial problem returning: relative to all that, if U is always unaccelerated, it is as possible that U has one mass as that it has another.

So the necessity cannot be understood as 'nomological' in the sense of 'deriving from laws plus kinematic factors'. It would have to be a sort of necessity that is specific and different from body to body. In other words, this program needs very specific de re modalities or essences or the like. Suppose, however, that the mode of response of body U to a certain kind of experiment is introduced as an essential property of U, by postulate. Then we can hardly count the manner in which mass has been eliminated from the primitive concepts of mechanics as a gain over the reliance on counterfactuals.

1.1.3 Alternative approaches to mass in mechanics

In the axiomatic theories of mechanics developed in this century, we see many different treatments of mass. In the theory of McKinsey, Sugar, and Suppes (1953), as I think in Newton's own, each body has a mass. In

Hermes's theory, the mass ratio is so defined that if a given body never collides with another one, there is no number which is the ratio of its mass to that of any other given body. In Simon's, if a body X is never accelerated, the term 'the mass of X' is not defined. In Mackey's any two bodies which are never accelerated, are arbitrarily assigned the same mass.[2]

What explains this divergence, and the conviction of these authors that they have axiomatized classical mechanics? Well, the theories they developed are demonstrably empirically equivalent in exactly the sense that *any phenomena which can be accommodated by a model* of any one of them can be thus accommodated by all. Therefore, from the point of view of *empirical adequacy*, they are indeed equal. And this, an empiricist would wish to submit, is just the basic criterion of success in science, to which all other criteria are subordinate.

1.2   Attempted eliminations of theoretical terms

The dispute about mass is one specific example of the wider problem of how to understand *theoretical terms*, that is, terms newly introduced to formulate new theories that could apparently not be expressed in the language up till that moment. While the program of providing defnitions of the sort that Mach sought, or Bridgman's 'operational definitions' for all such terms, or the early positivist attempt to understand the language of science through its relation to just 'observation vocabulary' alone, have long since been definitively rejected, there is still a related idea. That is the persistently seductive philosophical notion that all theoretical terms have a meaning precisely determined by the roles they play in scientific discourse. That is the idea of *implicit definition*.

Hilbert is generally credited with making this idea precise. When it ran into heavy weather, many logical tours-de-force were tried to either defend or reformulate it in defensible form. After Hilbert, the idea journeyed through writings of Frank Ramsey, David Lewis, and Frank Jackson to morph recently into the 'Canberra Plan' (Braddon-Mitchell and Nola 2009b). Here I will describe some of this history, and the argument that new developments in logic - most especially, Beth's theorem on definability – should have spelled the end of these attempts to use the notion of 'implicit definition' and rescued philosophy of science from the seduction of this mirage.

1.2.1   Hilbert's introduction of 'implicit definition'

Early in the century, Hilbert introduced the notion of "implicit definition" in connection with the meanings of terms in geometry. In commenting on his own axiomatization of Euclidean geometry, where such terms as "point" and "between" occur as primitive, Hilbert wrote "The axioms of [order] define the idea expressed by the word 'between'." (Hilbert 1902: 5) And more generally, he took the axioms to be components of the definition of the terms that are primitives of the theory.

What are we to make of this? At first blush, Hilbert's proposal may sound very plausible. When we understand the axioms, and are able to deduce theorems, to solve problems posed concerning the system, how could we be said not to understand what we are doing? Yet the doing consists entirely in the systematic use of the terms introduced in the formulation of that theory ... . But the question started with the words "When we understand the axioms"; how could we be said to understand the axioms if we do not already understand the terms in them?

Hilbert's views about "implicit definition" were immediately subjected to criticism by Frege in correspondence. Hilbert did not agree to have the correspondence published, but Frege then presented his side in a review (Frege 1903), to express what must surely puzzle everyone about Hilbert's notion. Perhaps the word "between" will have its meaning fixed by some axioms, in which other terms occur that we already understand. And perhaps another of those terms could have its meaning fixed by those axioms if we take "between" as understood. But circularity seems to threaten if we suppose that **all** the terms occurring in the axioms have their meaning fixed in this manner.

I say "seems to threaten"; not everyone has seen this as a real threat. In fact, despite Frege's vigorous critique, the idea proved tremendously appealing and continued in a long life, though in various forms, often highly ingenious and always apparently responsive to philosophical puzzles.

1.2.2   Quine's 'Implicit definition sustained'

In an article that we must mainly read as ironic, Quine [1964] purported to have rescued the idea of implicit definition, to his own dismay. His dismay with the idea is clearly expressed at the beginning:

> What is exasperating about the doctrine is its facility, or cheapness, as a way of endowing statements with the security of analytic truths without ever having to show that they follow

from definitions properly so called, definitions with eliminable definienda. (Quine 1964: 71)

So what is the argument that purports to support that doctrine? Assume that a certain empirical theory -for example, chemistry- is true, and can be formulated with predicates $F_1, ..., F_n$ and a single axiom $A(F_1, ..., F_n)$. Then it is satisfiable, and there will be a structurally similar statement in arithmetic $A(K_1, ..., K_n)$ that is an arithmetic truth. (Here Quine is drawing on the Loewenheim-Skolem theorem and some related results in metalogic; we'll return to those in a moment.) Now proceed as follows: have a language that includes arithmetic and also predicates $G_1, ..., G_n$ which are interpreted to mean the same as the original chemical predicates $F_1, ..., F_n$. Do not introduce any axiom at all! As Quine shows, it is now possible to define the predicates $F_1, ..., F_n$ in terms of those new predicates G_i and defined predicates K_i so that

1. the statement $A(F_1, ..., F_n)$ will be true just because $A(G_1, ..., G_n)$ happens to be true (assumption!) and $A(K_1, ..., K_n)$ is an arithmetic truth, and

2. the statement $A(F_1, ..., F_n)$ is deducible from the arithmetic truth $A(K_1, ..., K_n)$

How did Quine perform this leger-de-main? The assumption that the theory he is considering is true played a crucial part. On that assumption, the augmenting definitions do indeed introduce expressions co-extensive with what they purport to define; but only on that assumption (cf. Wilson 1965).[3] So the theory is 'mimicked' among the arithmetic truths; but the idea that the formula $A(F_1, ..., F_n)$ is sufficient to give meaning to the predicates $F_1, ..., F_n$ is spurious. Quine himself drops his ironic tone toward the end, and calls the manoever "farcical" and "hocus pocus".

1.2.3 Keeping something fixed: Winnie's rejection

While Quine kept something - the truth of the empirical theory in question – fixed rather surreptitiously, Winnie [1967] proposes that we explicitly suppose the extension of 'observational' terms to be fixed, and asks whether the remaining 'theoretical' terms can then be said to be implicitly defined by the axioms of the theory. The conclusion he reaches is negative.

Winnie assumes that the domain of discourse is divided into two disjoint parts. The candidates for referents of these two sorts are thought of as the

'theoretical' and the 'observational'. If we look at a model of the theory, we see that certain permutations of the 'theoretical' part need not affect the truth-value of any of the theorems. So the extension of the theoretical terms is not fixed by the theory, even on the assumption of fixed values for the observational terms. Therefore the theoretical terms cannot be said, in any sense, to be 'implicitly' defined by the role they play in the theory.

1.2.4   Lewis on the definition of theoretical terms

Lewis [1970] remarked that in the case of Winnie, there was still a trick involved: it was not a surreptitious fixing, but rather variation, that played the tricky role. Some predicates have extensions bridging the two parts of the domain, although they are among old terms, that the scientists had before formulating the new theory. For example, before atomic theory developed in that direction, "larger" was an established term. But in the theory this same term is also utilized (once the new terms are available) in such theoretical assertions as that salt molecules are larger than sodium atoms. In Winnie's reconstructions the extensions of those terms are allowed to vary 'on one side', so to speak, but not on the other. Their application on the 'observational' side remains the same, but in other respects, their application is taken to be up for grabs.[4]

In his own proposal, Lewis insists that the new terms will receive explicit definitions, relative to an assumption of fixity for all the old terms. He undertakes to reconstruct the theory so that this will be possible. The first stage in arriving at the 'correct' formulation of the theory is to replace it by its Ramsey Sentence – that is, a second order formula in which the theoretical predicates are replaced by bound predicate variables. It is easiest to explain what a theory's Ramsey sentence is by giving a toy example. The little theory

> Water consists of hydrogen atoms and oxygen atoms.

has as its Ramsey sentence:

> There exist three properties such that water is composed of entities which have the first and third property and entities which have the second and third property.

A caricature of an example, of course. But the crucial point is the same as for any more extensive theory: all consequences of the original theory which can be stated entirely in the old vocabulary are also consequences of its Ramsey sentence.

However, if left there, Winnie's point could be made: this Ramsey sentence surely could be true in many ways. Therefore the Ramsey sentence does not fix the meanings of the terms which we wanted to understand, the ones that are noticeable precisely for their absence in that sentence! But Lewis's construal of the theory has a second stage: the Ramsey sentence is its first postulate, but as second postulate it has an assertion to the effect that this Ramsey sentence has a unique instantiation.

What form can this second stage really take? Here Lewis steps in Carnap's footsteps, adding the *Carnap Sentence*. Whereas the Ramsey sentence says in effect that the theory has a realization, the Carnap sentence says that if it has a realization, then that realization is unique. We may call their conjunction the *Lewis Sentence*. Among the sentences entirely formulated in old terms, the Carnap sentence implies nothing at all. The two postulates together imply only what the original theory implied. To illustrate: for our little 'water theory' the complete new formulation of the theory would amount to:

> *Lewis Sentence*. There exists three **and only three** properties such that water is composed of entities which have the first and third property and entities which have the second and third property.

It is true that, with a little ingenuity utilizing definite descriptions in second-order language, this condition sanctions explicit definitions of those properties.

There is however a problem that defeats Lewis's attempt. Lewis initially understood the properties mentioned in the Ramsey sentence extensionally: two properties are the same exactly if they have the same instances. In effect, the quantifier "there are properties" is read as ranging over sets. But a famous argument by Hilary Putnam (1978), not very long after Lewis's article, shows that the Ramsey sentence must then have many different ways of being true. So the Lewis sentence – which asserts unique instantiation – is necessarily false!

That this is so, no matter how large or complex the theory, Putnam showed on the basis of the same sorts of meta-logical results that were mobilized earlier by Quine and Winnie: Loewenheim and Skolem's famous theorems and their later more powerful extensions (cf. van Fraassen 1997). These theorems show that if a theory has an infinite model then it has models of every infinite size; of course, models of different sizes do not have the same structure.

When Lewis himself realized this, perhaps because of Putnam's argument, he introduced a further postulate, to the effect that some sets mark 'natural' divisions in nature, and some are merely 'arbitrary'. (That there is such a division in nature, that our predicates must "carve nature at the joints", is a standard idea in the sort of metaphysical realism that Putnam was attacking.) With that metaphysical postulate in place, the Lewis sentence can be modified to assert, in effect, that the Ramsey sentence is uniquely instantiated among the *natural sets*.

This modification of Lewis's view cannot be refuted by Putnam's argument, but only because there is no information about which sets are natural and which not. In any case, the contention that the new terms can be explicitly defined can now be said to be correct only relative to this postulate about a division in nature between natural and arbitrary divisions. That is really a far cry from the original idea that the new terms are wholly understood given our understanding of the old terms.

1.2.5   The relevance of Beth's theorem

As we just saw, Lewis appreciated that there is not such a great distance between implicit and explicit definition. For him this opened the hope that new theoretical terms could after all be explicitly defined, and hence understood, in old terms. But the results in meta-logic mentioned above, point in the opposite direction: that the impossibility of explicit definition also eliminates any genuine or non-trivial sense of implicit definability.

This is where the philosophy of science could and should have benefitted clearly from attention to Beth's theorem (Beth 1953). First of all, the notions of explicit and implicit definability are there clarified to the point of being equipped with precise, applicable criteria. The notion of explicit definition itself is not without its complexity (cf. Wilson 1965). Suppose we write, at the beginning of a theory, for a certain predicate X,

$$X(x) = def Y(x)$$

Then we can use the sentence $(x)(Xx \equiv Yx)$ in any proof, just in the way we can use a tautology. But the former is a statement outside and about the theory, whereas the latter is a statement in the language of the theory. There are severe restrictions on this practice, despite the popular impression that we can always just define words any way we like. For example, we cannot introduce the new name a by defining the predicate

$$a = x = def Y(x)$$

unless it is part of the theory that the condition $Y(...)$ is uniquely satisfied. For example, we can't define the number k by the equation of "$k = x$" with "$x = x.x$", because $1 = 1.1$ but also $0 = 0.0$, so that would have the effect of implying that $1 = 0$.

So Beth rightly decided that the two notions "explicitly definable in theory T" and "implicitly defined by theory T" need criteria of application. The criteria he offered do not go back and forth between what is done 'outside' the theory and what is admissible in deductions 'inside' the theory, thus avoiding such difficulties as the above. For a one-place predicate F, and a set of axioms $\mathbf{A}$, we have the criteria:

> F is explicitly definable relative to $\mathbf{A}$ in terms of $G_1, ...G_n$ if and only if there is a formula $U(G_1, ..., G_n)$ in which no other non-logical signs occur, such that
>
> $$(x)[F(x) \equiv U(G1, ..., Gn)(x)]$$
>
> is derivable from $\mathbf{A}$
>
> F is implicitly defined by axioms $\mathbf{A}$ if and only if, for a predicate G not occurring in $\mathbf{A}$, the sentence $(x)[F(x) \equiv G(x)]$ is derivable from the union of $\mathbf{A}$ and $(G/F)\mathbf{A}$

where (G/F) is the operation of replacing all occurrences of F by occurrences of G. With these definitions in hand, we have the result which ends once and for all the suggestion that there is a non-trivial extended sense of "definition" in which a theory defines, fixes the meaning, of its own primitive terms:

> **Beth's Theorem**. The predicate F is implicitly defined by axioms $\mathbf{A}$ if and only if F is explicitly definable relative to $\mathbf{A}$ in terms other than F.

Between the results of Loewenheim and Skolem and Beth's theorem the idea that theoretical terms cannot be explicitly defined, yet have their meaning presented implicitly within a pre-theoretical understanding, by their roles in inference, loses any basis it could possibly have.

2   The Problem of Old Evidence

Traditional epistemology, focused on warrant, justification, and knowledge, flourishes still today, but there is little contact between its literature

and that of the philosophy of science. Yet epistemological concepts such as *evidence*, and a concern with the rationality of changes in our view of what the world is like, are central in philosophy of science as well. One development that has separated the two areas in our discipline is the influence of *probabilism* in the latter. That is the view that the form of opinion is best understood as judgment made in terms of probability. This includes certainty (probability equal to 1) as a limiting case, not at all precluded, but to be treated within the continuum of degrees of probability.

At the same time, it is admitted everywhere that the probability calculus is just a beginning, and that there are serious problems remaining for the modeling of opinion and rational opinion change. The *problem of old evidence* is one of these problems; it concerns the rationality of opinion change when the reason given is some information that was actually already known for some time before the change. For example, the facts about the perihelion of Mercury had been known for decades, but were cited as evidence for Einstein's General Theory of Relativity when that was proposed. If the prior opinion was rational and coherent, while that old evidence was already known, how could it now suddenly be brought up to demand a change in view?

2.1   Evidence, confirmation, and probability

A simple, perhaps the most naïve, form of probabilism is known as the orthodox Bayesian position, which includes a strict rule for how opinion must accommodate new evidence. This picture of our epistemic situation has evidence coming to the agent/subject in the form of a proposition E. Taking E as evidence consists in amending the prior opinion by *simple conditionalization* on E, which can be explained as follows.

The prior opinion is represented by a probability function P, which assigns probabilities to all the propositions the agent understands. The number $P(A)$ that P assigns to proposition A is called the agent's prior credence in A. The new, posterior, opinion is then represented by the function P_E – read as "P conditionalized on E" – which is defined by:

> If A is any proposition (for which the prior credence is not zero) then the posterior probability $P_E(A)$ is just the ratio of two prior probabilities, namely $P(E$ and $A)$ divided by $P(E)$.

(We also write $P(A|E)$ for $P_E(A)$. When $P(E) = 0$, the conditional probability is not defined.) The concept of *confirmation* typically associated with this simple account is this:

> proposition E *confirms* hypothesis H (itself a proposition) for an agent if his posterior credence in H – his credence in H conditional on E – is higher than his credence in H tout court.

There are various objections to this proposed explication of the concept of confirmation; we'll consider just one here. It was Clark Glymour [1980] who first emphasized the difficulty this engenders for evidence that had already been known. If the agent has already, previously, become certain that E, then $P(E) = 1$. From this it follows then that, in the above sense, E confirms no hypothesis whatever for that agent, for dividing by the number 1 has no effect.

That is difficult to reconcile with well known episodes of theory change in the sciences. It would make nonsense of such claims as that the advance in the perihelion of Mercury, known for a half-century before Einstein's work, confirmed Einstein's theory of relativity.

2.2   Responses to the problem of old evidence

Three possible responses come immediately to mind, at least in outline form. The first is that no scientist is ever truly certain about E, whatever it may be. That does not remove the problem on a practical level, for if the probability of E is negligibly smaller than 1 then dividing by it makes only a negligible difference. The second response is that if today we say that E, long since believed with certainty, confirms H, we make reference not to our present actual epistemic state but to some alternative(s) thereto – perhaps the opinion we had way back when, before learning that E. But that is not realistic, because what we learned since then is certainly playing a role in our present reasoning, which could hardly have been carried out a long time ago.

There is a third response that has promise, and this is what I will examine in detail. It is the response that in this sort of case, the phrase 'E confirms H' may well be used, but the attention is actually drawn not to E, but to something else (having to do with E) which is a new discovery, and does make H more probable. The idea is that, for example, when the scientific community got to the point when it could say that the advance in the perihelion of Mercury confirmed Einstein's theory, they had indeed just learned something new. The new information was not that fact about Mercury, but another fact that has to do with it. Conditionalizing on that new fact had increased their credence in Einstein's theory.

What could that mysterious new fact be? Daniel Garber proposed this response, and elaborated it (Garber [1982]).

2.3   Postulating new 'logical' evidence [5]

Put most bluntly, this response postulates that there is a special proposition, a function of H and E, which is newly learned at that point, and is such that conditionalizing on *that* proposition increases the probability of H.

Obviously Garber's real contribution can't just be this; his solution consisted in a proposed identification of that special proposition. The value and importance of Garber's work, and the reason why we speak here of a solution, is that he attempts to tell us what this proposition is – to identify a proposition which can play the required role and to show us that it can.

For Garber tells us that this proposition is to be identified as the proposition that H *implies* E, which he symbolized with the 'turnstile' symbol as $(H \vdash E)$. When an agent claims that previously known evidence E confirms H, he should be understood to assert that his posterior credence $P_{(H \vdash E)}(H)$ is greater than the prior credence $P(H)$. What scientists learned around the year 1919 is that Einstein's theory correctly predicted – that is, implied – the advance in the perihelion of Mercury, which was not implied by Newton's theory. It was learning this implication that confirmed Einstein's theory.

Clearly some larger story must be told about just what that special proposition is. The use of the word "implies" gives us a clue to what that story should be but hardly suffices by itself. In fact, our intuitive understanding of this word immediately suggests that the problem of old evidence will just return at second remove. If H implies E, is that not a logical truth, which this person must have all along assigned probability 1 on pain of incoherence?

Here Garber steps in to write a (neo-) Bayesian theory for the logically non-omniscient subject. For this subject, the proposition $(H \vdash E)$ which we read as 'H implies E', is logically contingent. Given our worries about this solution, we recognize that Garber should be sensitive to the demand that identification of the proposition called $(H \vdash E)$ as having that meaning be reasonably warranted. Indeed he is. In response, he imposes the validity of modus ponens as a minimal condition on the conceptual role of $(H \vdash E)$. That is, the meaning of this special proposition must be such that the inference from premises H and $(H \vdash E)$ to conclusion E is valid.

2.4   Probabilist versions of implication and modus ponens

To explain this, Garber needs a probability counterpart of the relation of implication. What he chooses, implicitly, is this: A implies B exactly if the conditional probability of B given A equals 1, for all probability functions, regardless of background assumptions. This bestowing of certainty, by the antecedent on the consequent, is meant to provide a probabilistic version of implication, that we can read intuitively as "Given the antecedent, the consequent is certain".

In view of the definition above of conditional probability, this can be stated equivalently as:

> A implies B if and only if for all probability functions P and all propositions K:
>
> 1. (i) $P(B|A \text{ and } K) = 1$ whenever $P(A \text{ and } K) > 0$, or
> 2. (i) $P(B \text{ and } A \text{ and } K) = P(A \text{ and } K)$

(This is a little redundant, since conditionalizing P on K, if $P(K) > 0$, produces another probability function.) When Garber imposes the validity of modus ponens as a condition on the language containing the statement $(H \vdash E)$ he does so with, in effect, the following condition:

> Condition K.   For all propositions K, the probability of E, given both H and $(H \vdash E)$ and K, if defined, equals 1.
>
> In symbols, $P(E|H \text{ and } (H \vdash E) \text{ and } K) = 1$ whenever defined;
>
> equivalently, $P(E \text{ and } H \text{ and } (H \vdash E) \text{ and } K) = P(H \text{ and } (H \vdash E) \text{ and } K)$

So far Garber's account. But the logician must ask: does the validity of modus ponens provide reasonable warrant for that identification? Conjunction and material implication obey modus ponens too: the inferences from premise H and either premise (H and E) or premise (not H or E) to conclusion E are valid too. These alternatives are entirely unsuitable. Conjunction and material implication would not do, because in the examples the agent does not learn (H and E), and he already knew that (not H or E) because he knew E.

The conclusion, apparently inescapable, is that Garber's solution has not been substantiated unless he can make it work in the presence of minimal adequacy conditions for the interpretation of $(H \vdash E)$ as 'H implies E', and that modus ponens alone is not enough.

2.5   The Conditional Proof Requirement

The most obvious extra condition to be demanded for $(H \vdash E)$ is the validity of some suitable form of conditional proof. That form of proof would be codified in the rule:

> if E is derivable from premises Z together with H, then $(H \vdash E)$ is derivable from Z

At this point we must proceed as sympathetically as possible, to see how Garber could have imposed such a requirement in addition to modus ponens without debilitating his solution.

As a first attempt we might try to bring in background knowledge that characterizes the moment of theory change. If H implies E for the agent at this moment, that may well be in part because the agent holds to be true some theory T, that perhaps he knows only by description as the theory axiomatized in some area of science, and does not pretend to understand it very well. The point is that the agent himself may not have insight into the content of background theory T, but learns the fact that T and H together imply E by the standards of classical logic. If that is the story, the characterization of the conceptual role of $(H \vdash E)$ is this:

> $(H \vdash E)$ expresses the logically weakest proposition which together with H and T implies E in classical logic.

That certainly satisfies both the rules of modus ponens and of conditional proof. But here logic steps in: it is unfortunately provable that this identifies $(H \vdash E)$ as the material implication [not $(T$ and $H)$ or E]. It will not do because at the prior time when we deemed T to be true and knew E, logical coherence required that we deemed $(H \vdash E)$ to be true as well, already. It could not appear as new evidence, logically contingent relative to the agent's prior opinion.[6]

Of course, similar puzzles have been encountered in the study of modal logic. So from a formal point of view at least we may see some other options. The sentence $(H \vdash E)$ will have to mean that H implies E in some more full-blooded sense. Perhaps it could be something like that H guarantees the truth of E in a large variety of cases and not only in the actual one – or something that bears some formal similarity to such modal statements.

So let us follow Garber's lead, in his reconstrual of the rule of modus ponens, and ask if we can impose some condition analogous to K to honor the intuition of the validity of conditional proof.

That means that we have to formulate conditions under which it is a matter of logic that $(H \vdash E)$ is certain, relative to given premises and background knowledge, whatever that may be. While the agent could perhaps newly learn that $(H \vdash E)$ in many ways, on the basis of logically contingent evidence that might be hard to describe, there must also be a broadly logical rule of 'conditional proof'. That is what we tell our beginning logic students for the "if ... then" explained as the material conditional, as well as later on for such variants as the physically or metaphysically necessary conditional.

So I would propose this: if it is clear that E is certain conditional on the supposition that H, (i) regardless of what our background knowledge is, and (ii) regardless of how we have needed to respond to new evidence, then $(H \vdash E)$ is already certain. Starting to put in the symbols for the relevant prior probability P and for how P might change in response to new evidence, we can rephrase this as:

> Condition KK If, for any possible posterior probability P' that could evolve from P, and any background knowledge K, it is the case that P'$(E|H \text{ and } K) = 1$, whenever defined, then $P((H \vdash E)) = 1$;

> Equivalently If, for any possible posterior probability P' that could evolve from P, and any background knowledge K, it is the case that P'$(E \text{ and } H \text{ and } K) = P$'$(H \text{ and } K)$, then $P((H \vdash E)) = 1$

But again triviality threatens. The question to worry about is what that set of probability functions is that comprises the possible posterior probability functions P' that could evolve from P in response to new experience. Logically speaking, at least, these could be all the functions that do not raise zeroes, that is, that assign zero to all propositions to which P assigns zero. Alternatively, they could be all the ways in which the current probability P could be conditionalized on new evidence. In either case, the role of $(H \vdash E)$ is indistinguishable from the material conditional: they receive the same probability always (cf. van Fraassen 1988: 161).[7]

So either Garber's new connective $\vdash$ cannot be read as "if ... then", or Condition KK is too strong, or the range of possible posterior probabilities P' must be especially restricted for that condition. Or perhaps this third approach to the problem of old evidence is also on a wrong road.

3   Conclusion

What conclusions should we draw from the examples we have now examined, of the relevance of logic to philosophy of science?

In the case of theoretical terms, implicit definability and its variants, we can conclude quite definitively that certain popular philosophical ideas do not survive confrontation with the Loewenheim-Skolem theorem and Beth's theorem.

In the case of the problem of old evidence, the conclusion is not quite so definitive. We can say this much: if there is a way to accommodate warrant provided by old evidence in the Bayesian recipe for rational opinion change, we certainly have not seen it. Exploration of what seemed the most promising approach came to naught. As mentioned, this problem is just one among many for the modeling of opinion in a probabilist framework. There is currently a great deal of on-going work in 'formal epistemology' to improve and extend such models of opinion and reasoning.

In both cases we took up an interesting and important problem. What we can conclude on a positive note is that attention to logic made it possible to formulate rigorous criteria of adequacy for any proposed solutions. In that way, it was possible to rule out certain solutions that were proposed, and thus clear the way for more nuanced and more sophisticated approaches to those problems. After that, it is up to us.

References

[Braddon-Mitchel & Nola 2009a] Braddon-Mitchell, David and Robert Nola (Eds.) [2009a] Conceptual Analysis and Philosophical Naturalism. Cambridge, MA: MIT Press.

insert "2009a"

[Braddon-Mitchel & Nola 2009b] Braddon-Mitchell, David and Robert Nola (Eds.) [2009b] "Introducing the Canberra Plan", pp. 1-20 in Braddon-Mitchell and Nola.

[Bressan 1973] Bressan, Aldo [1973] A General Interpreted Modal Calculus. New Haven: Yale University Press.

[Beth 1953] Beth, Evert Willem [1953] "On Padoa's method in the theory of definition", Indigationes Mathematicae 15: 330-339.

[Earman 1984] Earman, J. (Ed.) [1984] Testing Scientific Theories. Min-

nesota Studies in the Philosophy of Science X. Minneapolis: University of Minnesota Press.

[Frege 1903] Frege, Gottlob [1903] "Ueber die Grundlagen der Geometrie", Jahresbericht der deutschen Mathematiker- Vereinigung XII: 319-324 and 368-375; transl. "On the Foundations of Geometry," Philosophical Review, 69 (1960) : 3-17.

[Garber 1984] Garber, Daniel [1984] "Old Evidence and Logical Omniscience in Bayesian Confirmation Theory", pp. 99-132 in Earman 1984.

[Glymour 1980] Glymour, Clark [1980] Theory and Evidence. Princeton: Princeton University Press.

[Hermes 1959] Hermes, H. [1959] "Zur Axiomatisierung der Mechanik", p. 250 in Proceeding of the International Symposium on the Axiomatic Method held at Berkeley, 1957-58. North-Holland Pub. Co.

[Hermes 1938] Hermes, H. [1938] "Eine Axiomatiserung der Allgemeinen Mechanik", Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften, Vol. 3 (Heft 3) Leibzig: Verlag von Hirzel.

[Hermes 1959] Hermes, H. [1959] "Modal Operators in an Axiomatisation of Mechanics", Proceeding of the Colloque International sur la méthode axiomatique classique et moderne. Paris.

[Hilbert 1902] Hilbert, David [1902] The Foundations of Geometry. Chicago: Open Court.

[Jammer 1964] Jammer, Max [1964] Concepts of Mass. New York: Harper.

[Jeffrey 1984] Jeffrey, Richard C. [1984] "Bayesianism with a Human Face" pp. 133-156 in Earman1984.

[Lewis 1970] Lewis, David K. [1970] "How to Define Theoretical Terms", The Journal of Philosophy 67: 427-446.

[Mach 1883] Mach, Ernst [1883] The Science of Mechanics: a Critical and Historical Account of Its Development (Sixth English edition LaSalle, IL: Open Court Pub. Co. 1960).

[Mach 1896] Mach, Ernst [1896] Principles of the Theory of Heat. Tr. B. McGuinness. (Dordrecht: Reidel 1986).

[Mackey 1973] Mackey, G. W. [1973] The Mathematical Foundations of
    Quantum Mechanics. New York: Benjamin.

[McKinsey et al. 1953] McKinsey, J. C. C. , Sugar, A. C. , and Suppes, P.
    [1953] "Axiomatic Foundations of Classical Particle Mechanics", Jour-
    nal of Rational Mech. and Analysis 2: 253-272.

[Montague 1960] Montague, Richard [1960] Review of Simon 1959, Jour-
    nal of Symbolic Logic 25: 355-356.

[Painlevé 1922] Painlevé, P. [1922] Les axiomes de la mechanique. Paris:
    Gauthiér-Villars.

[Pendse 1937] Pendse, C. G. [1937] "A note on the definition and deter-
    mination of mass in Newtonian Mechanics", Philosophical Magazine
    Series 7 vol. 24: 1012-1022.

[Pendse 1939] Pendse, C. G. [1939] "A further note on the definition and
    determination of mass in Newtonian mechanics", Philosophical Maga-
    zine Series 7 vol. 27: 51-61.

[Pendse 1940] Pendse, C. G. [1940] "On mass and force in Newtonian me-
    chanicsaddendum to 'Mass I.' and 'Mass II'", Philosophical Magazine
    Series 7, vol. 29: 477 484.

[Putnam 1979] Putnam, Hilary [1978] "Realism and reason", pp. 123-140
    in his Meaning and the Moral Sciences. New York: Routledge.

[Quine 1964] Quine, Willard V. [1964] "Implicit Definition Sustained",
    The Journal of Philosophy 61: 71-74.

[Rosser 1938] Rosser, J. B. [1938] Review of H. Hermes, Journal of Sym-
    bolic Logic 3: 119-120.

[Simon 1947] Simon, Herbert [1947] 'The Axioms of Newtonian Me-
    chanics', Philosophical Magazine 38: 888-905.

[Simon 1954] Simon, Herbert [1954] "Discussion: The Axioms of Clas-
    sical Mechanics", Philosophy of Science 21: 340-343.

[Simon 1959] Simon, Herbert [1959] "Definable Terms and Primitives in
    Axiom Systems", pp. 443-453 in L. Henkin et al.(Eds.), The Axiomatic
    Method. Amsterdam: North-Holland.

[Simon 1966] Simon, Herbert [1966] "A Note on Almost-Everywhere Definability" (abstract), Journal of Symbolic Logic 31: 705-706.

[Suppes 1957] Suppes, Patrick [1957] Introduction to Logic. Princeton: Van Nostrand.

[van Fraassen 1988] van Fraassen, Bas C. [1988] "The problem of old evidence" ", pp. 153-165 in David F. Austin (ed.) Philosophical Analysis. Dordrecht: Reidel.

[van Fraassen 1997] van Fraassen, Bas C. [1997] "Putnam's Paradox: Metaphysical Realism Revamped and Evaded", Philosophical Perspectives, vol. 11: 17-42 (Boston: Blackwell).

[Wilson 1965] Wilson, Fred [1965] "Implicit Definition Once Again", The Journal of Philosophy 62: 364-374.

[Winnie 1967] Winnie, John A. [1967] "The Implicit Definition of Theoretical Terms", The British Journal for the Philosophy of Science 18: 223 -229.

Notes

[1]Montague (1960) is a very critical, quite dismissive review of Simon 1959. Simon's idea is in itself easy enough to understand, however, and the shortcomings of his formal treatment are only amusing.

[2]See Jammer, op. cit., Ch. 9; Mackey, pp. 14; compare Simon's first approach in his 1947 and 1954.

[3]Quine had an ulterior motive for writing this ironic article. The lesson he wants us to draw is clear: philosophers he opposed had taken chemistry to be synthetic and arithmetic analytic. When the true and supposedly synthetic assertions in chemistry can be deduced from the supposedly analytic truths of arithmetic augmented with definitions, this analytic/synthetic distinction is rendered vacuous.

[4]Paul Feyerabend is well known for the contention that when accepted scientific theory changes, the meanings of the old terms, used in observation reports, does not remain the same either. While Lewis may be read here as objecting that Winnie assumes a special stability for the old terms, so does Lewis.

[5]For a full account see Jeffrey 1984 , Earman 1984, van Fraassen 1988.

[6]Note especially that this reasoning does not require that the auxiliary theory T be expressible in the object language – only some meta-linguistic understanding is required on the part of the speaker to become clear on what sort of thing he means by this turnstile.

[7]The validity of modus ponens, Condition K, already shows that $P(\sim E \& H \& (H \vdash E)) = 0$ so that $P(\sim (H \& \sim E)|(H \vdash E)) = 1$. More importantly, consider the function P* which is P conditionalized on the material conditional $\sim (H \& \sim E)$. Clearly, for any

K, $P * (E\&H\&K) = P * (H\&K)$, so by Condition KK, $P * ((H \vdash E)) = 1$. But that means that $P((H \vdash E)| \sim (H\& \sim E)) = 1$.