Henry E. Kyburg, Jr.

Theories as Mere Conventions

1. Expunging Conventional Elements.

Conventionalism, as an approach to understanding scientific theory, is hardly new. Mach[1] construed the Newtonian relation $f = ma$ as a conventional characterization of the unobservable quantity "force." Poincaré[2] argued that the geometrical structure we attribute to space is a matter of convention. These two examples are interestingly different. Poincaré writes as a mathematician: whatever the true laws of dynamics may be (provided they satisfy very general and uncontroversial conditions) they can be expressed in terms of either Euclidean or non-Euclidean space. Mach writes as an experimentalist: the only fundamental way you can tell about forces is to measure masses, times, and distances–i.e., masses and accelerations. Yet these widely different points of view lead Mach and Poincaré to the same general view regarding certain theoretical claims.

To Duhem[3] is attributed the thesis that *theoretical* as opposed to *experimental* laws are essentially conventional and irrefutable. The link between irrefutability and conventionality is firmly established in Duhem. And the modern representative of this view, though in somewhat attenuated form, is Quine.[4] Quine's thesis, though bracketed with Duhem's, is somewhat different, for Quine sees the poles, "conventional" and "substantive," as unrealized ideals. That is, every statement in our scientific corpus partakes of both analyticity and empiricalness. The distinction is one of degree: some statements are more resistant to counterevidence than others.

Quine's view will be our starting point. As metaphored by Quine and Ullian,[5] our body of scientific knowledge should be viewed as a web. At the center of the web are the truths of logic and mathematics–the statements most resistant to modification in the face of recalcitrant experience. At the periphery of the web are statements most directly related to experience, to what happens to us. These are the occasion sentences on which our whole body of scientific knowledge is to be based. At the center, then, according to this picture, we have sentences that are highly structured and regimented, but evidentially remote from our ordinary

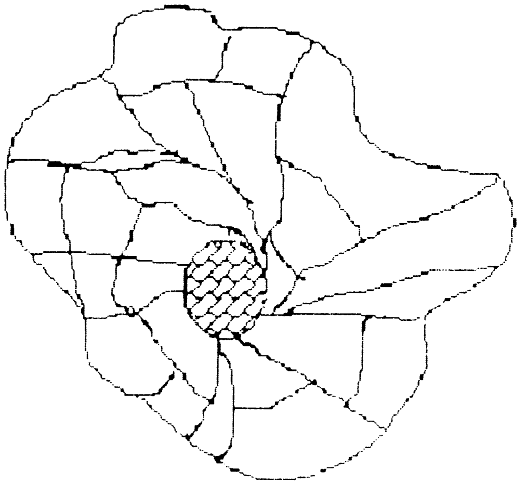

Figure 1.

experiences. At the edges, we have sentences that are more or less directly warranted by ordinary experience, but that have little or no structural connection to the rest of our body of beliefs. The general picture is reflected in figure 1.

Let us now consider how this view of things would be changed if we were to drop from our representation the arbitrary elements. That is, if logic is thought

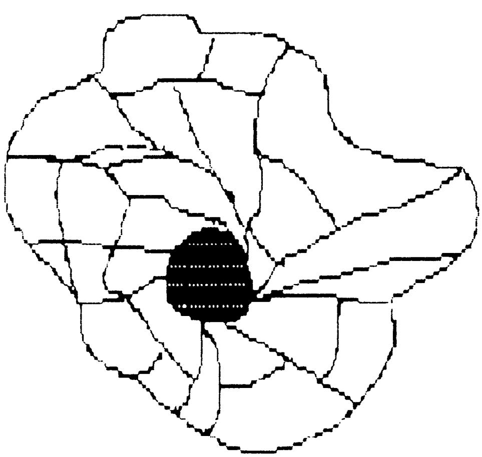

Figure 2.

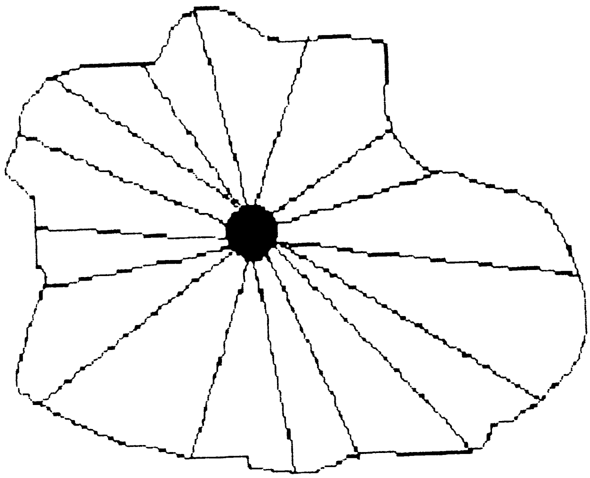

Figure 3.

of as being arbitrary or conventional, let us delete from our picture the elements that correspond to logic. The view we have then is illustrated in figure 2. There is still a web, embodying many interconnections, on which experience impinges only at the periphery. We may think of the internal nodes of this web as representing "theoretical" terms, while the nodes at the edges represent "observation" terms.

But then, as Craig[6] showed us, the theoretical terms are arbitrary and inessential additions—conventional elements that, but for convenience, could be eliminated. If we follow this path, we arrive at the prettier web illustrated in figure 3, in which there is a central node (the axioms of the physical theory, expressed in purely observational terms) connected by deductive spokes to the observational consequences on the rim.

If this picture of knowledge strikes you as *too* simple, you can, as I once pointed out,[7] impose your own favorite structure, involving your own favorite theoretical entities. More explicitly: you can start with a theory $\mathbf{T}_1$, choose your favorite theoretical entities, your favorite (consistent) laws concerning these entities, and be sure that there exists a theory $\mathbf{T}_2$ that (a) employs as its only theoretical entities those you have chosen, (b) has as its theoretical laws those you have chosen, (c) has exactly the same observational consequences as $\mathbf{T}_1$, and (d) requires those laws for the deduction of the observational consequences.

It appears that only the edge of the web is nonconventional. If we eliminate all the conventional elements we are left with figure 4.

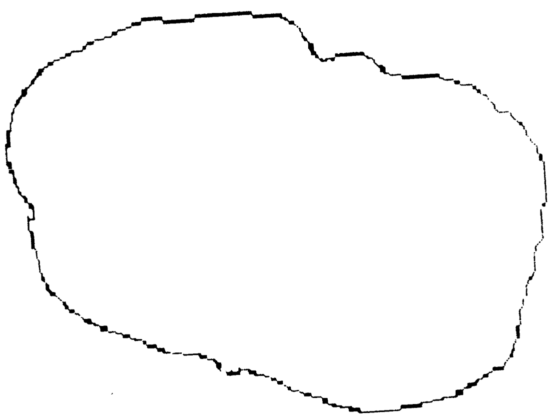

Figure 4.

All this has been predicated on a very shakey assumption—namely that the edge of the web, the observation sentences, can be sharply distinguished from the interior of the web, sentences involving theoretical terms. But to reject that distinction leads to an even worse state. If we associate irrefutability with convention, as we have been doing, then even the observation sentences begin to seem conventional. As Quine pointed out long ago,[8] even observation sentences are not uniquely determined by what happens to us. Alternatively, we may reflect that *any* observation sentence, if it has content at all, may turn out to be in error.

If we give up the distinction between theoretical and observational vocabulary, and disregard the conventional elements, we are left with a web of knowledge that has neither interior nor periphery.

This is clearly unproductive; we must somehow escape this consequence. The way to do so is to take a closer look at conventionality and the way in which it functions in scientific knowledge. Up to this point I have been using the term "convention" in what I take to be an ordinary philosophical sense. This is the sense discussed by David Lewis,[9] in which arbitrariness is the essence of convention. Indeed, Lewis writes "It is redundant to speak of an *arbitrary* convention. Any convention is arbitrary because there is an alternative regularity that could have been our convention instead" (p. 70).

To say this immediately sets lights flashing. Poincaré certainly did not regard Euclidean geometry as conventional in *this* sense! In fact, he (mistakenly) argued that Euclidean geometry would always be the geometry of choice for physical theory, on the ground of its simplicity. This is exactly to claim (a) that the choice of a geometry is conventional, and (b) that there are reasons for choosing one convention over another. If there are reasons for choosing one convention over another, these conventions are not "arbitrary." It is hard to believe that any conventions are entirely arbitrary; "arbitrary convention" seems to me more like an oxymoron than a redundancy. But be that as it may, we may certainly single out a class of conventions that may be adopted for *epistemic* as opposed to *nonepistemic* reasons.

Conventions that are adopted for nonepistemic reasons, we shall call *arbitrary*. Conventions that are adopted for epistemic reasons, we will not call arbitrary, even though "there is an alternative regularity that could have been our convention instead."

And what is a "nonepistemic" reason? Familiarity, elegance, and the like. I shall also construe simplicity as nonepistemic, though, as is well known, some philosophers take the simplicity of a scientific theory to be grounds for taking it to be acceptable, or true, or both. Epistemic reasons I take to be reasons expressible in terms of predictive power or probability. They will become clearer shortly. The general thesis is that scientific theories are *mere* conventions, but not *arbitrary* conventions. The reasons for choosing one theory rather than another are to be epistemic in my narrow sense.

2. Generalization, Observation, and Error

We consider a set of alternative languages, **L**, **L**′, etc. Each language contains a set of

(a) terms,
(b) formulas,
(c) axioms, and
(d) rules of inference

in the usual way. We suppose that

(i) First-order logic and set theory are included.
(ii) "Logical" or "analytic" relations among terms are derivable.
(iii) Probability as well as provability, as relations between sentences and sets of sentences, is defined.
(iv) An inductive as well as a deductive logic is provided.

With regard to (i), the intent is merely to have on hand as much mathematics as we need to do statistics and as is involved in whatever theories we are concerned with.

With regard to (ii), the intent is merely to capture those truths, if any, that are to be regarded as characteristic of the language in question. It is the burden of the third section of this paper to elucidate this vague characterization.

With regard to (iii), we must be a little more explicit. Probability is not taken to be subjective. It is *not* to be understood as a Carnapian degree of confirmation.

It is *not* (as both of these views would have it) based on a measure defined on the sentences of the language.

Probability is:

(1) defined for a given language **L**,
(2) relativised to a corpus of knowledge **K**,

where **K** is understood to be a set of sentences in **L**,

(3) syntactically definable,
(4) based on the statistical syllogism, and
(5) interval valued.

The statistical syllogism has the form:

Between 25% and 30% of *A*'s are *B*'s.
Charlie is an *A*.
That's all that is known in **K** that counts.
Therefore the probability of "Charlie is a **B**," relative to **K**, is [.25,.30].

The trick to making this notion of probability fly, of course, is spelling out in a noncircular way the condition embodied in the third premise.[10] I will simply assume, for present purposes, that it can be done. The upshot is that probability assertions have the form

$$\mathbf{Prob}_{\mathrm{L}}(S,\mathbf{K}) = [p,q],$$

where **L** is a specified language, *S* a sentence in that language, and **K** is an actual or hypothetical set of statements of that language. $[p,q]$ is just a closed subinterval of [0,1].

With regard to (iv), I assume, what is somewhat controversial, that we can come to accept statements on the basis of strong but inconclusive evidence. This is what I take inductive logic to be about. Just as with probability, I shall simply assume for present purposes that it makes good sense to talk about accepting statements on the basis of strnog but inconclusive evidence. I shall furthermore assume that we can operate with a purely probabilistic rule of acceptance.

Roughly the rule has the following form: if the probability of *S*, relative to **K**, is at least $1 - \varepsilon$, accept *S*. *More exactly, however, we must specify* **K**, and we must say *where S* is to be accepted. (If we look on logic as atemporal, we can't have *S* belong to **K**, for its probability would already be [1,1], which is not our intent.)

It turns out that a useful structure for present purposes can be represented by three levels of rational corpora. We take **K*** to be the corpus of incorrigibilia: it comprises logical and mathematical truths, analytic statements, and possibly incorrigible observation statements. The next level of corpus, **K**, is the evidential corpus. It contains evidential certainties, indexed by a real number p. A statement is acceptable as an evidential certainty if its probability, relative to the set of incorrigibilia, is at least p. Finally, the corpus on the basis of which we act, and relative to which the probabilities of the various possible outcomes of our actions are defined, is the corpus of practical certainties, $\mathbf{K}'$. This is indexed by a real number p' smaller than p. A sentence S is acceptable as practically certain if its minimum probability, relative to the evidential corpus, exceeds p'.

We thus have two parameters, p and p', to account for. I think a sensible story can be given, but this is not the place for it.[11]

We need just a little statistical inference. I claim, without argument here, that we can inductively infer a proportion in a population from the observation of a proportion in a sample. This makes use of a set theoretical truth:

> For perfectly reasonable n, ε, and δ, the proportion of n-membered subsets of a finite set A that exhibit a proportion of Bs that is within ε of the proportion of Bs among As in general is at least $1 - \delta$.

The argument also depends on the fact that the "third condition" of the statistical syllogism may be satisfied with regard to a given subset of A, relative to plausible bodies of knowledge **K**.

From these premises, we obtain: The probability that this particular sample, having its given proportion of Bs, exhibits a proportion that is close to that among As in general, is $[1 - \delta, 1.0]$.

Assume that our counting is incorrigible, so that the data is in $\mathbf{K}^*$, and that $1 - \delta$ is greater than p, and we have shown that "The proportion in our sample is r and the proportion in A differs from that in the sample by less than ε" is entitled to be in our corpus **K** of evidential certainties. Since this statement entails "The proportion of As that are Bs is in the interval $[r - \varepsilon, r + \varepsilon]$" the latter statement, too, may be in **K**.

We need one more piece of machinery to examine the possibility that scientific theories may be construed as mere, but not arbitrary, conventions. We must consider also, corresponding to each language **L**, its metalanguage **ML**. In each case **ML** contains **L** as a sublanguage. In addition, **ML** is assumed to contain a single nonlogical primitive relation, $\mathbf{O}(X, S, t)$, which is to be interpreted thus: the individual X, our agent, an individual, society, group, or whatever, *observes* the state of affairs or event denoted by the sentence S of **L**, at some time earlier than t.

With the help of this metalinguistic predicate, we can define two important classes of sentences of **L**:

D1 $VO(X) = \{S \mid (\exists t)(\mathbf{O}(X, S, t) \,\&\, S)\}$

These are X's veridical observations.

D2 $EO(X) = \{S \mid (\exists t)(\mathbf{O}(X, S, t) \,\&\, S)\}$

These are X's erroneous observations.

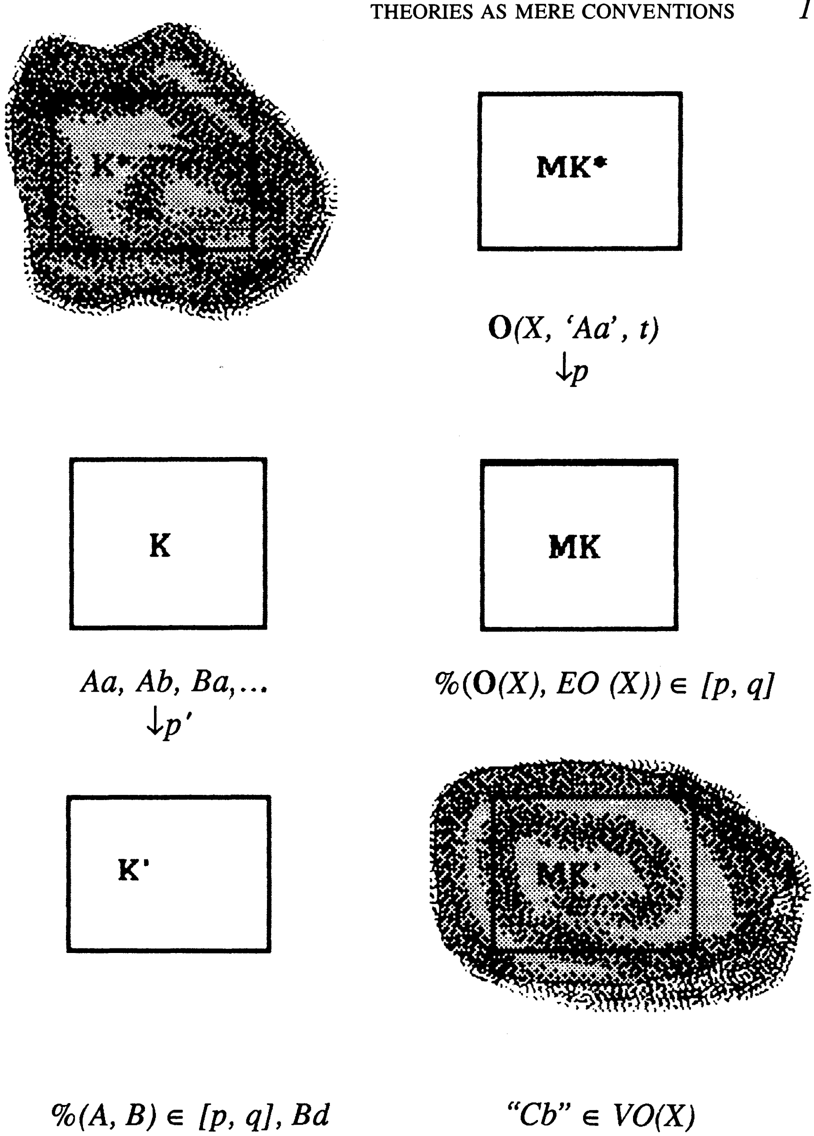

Figure 5.

Finally, we add one principle to our inductive logic: that if a statement in **L** has a probability greater than p relative to the evidential metacorpus, then it may appear in the object language corpus of evidential certainties.

Figure 5 may help to give the general idea of the structure I have in mind. The corpus of incorrigibilia in the language **L** seems to serve no useful purpose; logical truths and the like will enter the corpus of evidential certainty by being highly probable (having probability 1) relative to the metacorpus of evidential certainties. Observation statements themselves *may* appear in **K** when they are the result of observation: what is required is that they not be known to belong to any class of observations in which error is rampant. (If the observation is made under bad light, it is more likely to be erroneous—and we can come to know this in our metacorpus.)

For an example of how this works, let us consider the familiar ravens. This will also constitute an initial example of my main point. Consider two languages, $\mathbf{L}_1$ and $\mathbf{L}_2$, that both contain the predicates "is a raven" and "is black." $\mathbf{L}_2$ contains *as a meaning postulate* the generalization that all ravens are black. This is construed as a priori, as a *linguistic* constraint, embodied in the language itself.

Let us now consider the question, "Given a set of observation reports, how do we choose between the two languages?" (This reflects Mary Hesse's distinction[12] between observation *statements* and observation *reports*. The observation report is the linguistic entity entered into the laboratory notebook. The observation statement represents the corresponding fact. It is bad form to correct an observation report; but observation statements may sometimes turn out to be wrong and need to be corrected.)

Suppose our agent X has accumulated a long list of observation reports in his incorrigible metacorpus; statements such as $\mathbf{O}(X$, "a_1 is a raven", $t_4)$. These statements may be the same in either language. In particular, there is no inconsistency, even in $\mathbf{ML}_2$, between

$\mathbf{O}(X$, "a_1 is a raven", $t_4)$, and

$\mathbf{O}(X$, "a_1 is not black", $t_4)$,

though of course the two mentioned statements are inconsistent in both $\mathbf{L}_2$ and $\mathbf{ML}_2$. In either case, our agent is simply reporting his observations.

The difference is that in $\mathbf{L}_1$ we have no reason to doubt any of the mentioned statements. It seems perfectly natural for the mentioned statements to pass directly into the body of evidential certainties $\mathbf{K}_1$. Given the vagaries of human observation, we may suppose that even under the best of circumstances, not all observed ravens will have been reported to be black. Most will be, and relative to the evidence embodied in $\mathbf{K}_1$, it may become practically certain, in $\mathbf{K}_1'$ that almost all ravens are black; that is, "%(is a raven, is black) ε [.9,1.0]" may be a member of the practical corpus. (Even if, unnaturally, *all* the observed ravens have been observed to be black, this could be the strongest statement acceptable on the basis of probability in the first language.)

This would allow our agent to be pretty sure that the next raven he saw would be black, or that a random raven would be black. It does not, in itself, provide any grounds for supposing that a nonblack thing is a nonraven.

Consider the other language, in which it is a priori true that all ravens are black. Given the same set of observation reports, including some reports of non-black ravens, the situation of our agent X is quite different. Now he *knows*, on syntactical grounds alone, that some of his observation reports are erroneous. Some observation statements, corresponding to observation reports, must be rejected. Which ones? How many?

We are free, of course, whichever language we adopt, to suppose that all or almost all of our observations are erroneous. This is silly skepticism; we have no grounds for making any such assumption. So I propose that we adopt the following principle:

> $\mathbf{P}_1$ *The minimization principle:* Attribute no more error to your observations than your language requires you to.

This principle tells us how many observations we are required to regard as erroneous, but it doesn't specify which ones. It is not unreasonable not to be able to tell whether it is the observation report "a is a raven" or the observation report "a is not black" that is wrong. We could reject them both, but that would fly in the face of the minimization principle. The reason we want to know how many observations of each kind are in error, is that we want a statistical basis for determining long-run error rates for the various kinds of observations. A natural principle to follow (though it is not as compelling as the previous one) is this:

> $\mathbf{P}_2$ *The distribution principle:* Distribute the error frequencies as evenly as possible among the (syntactic) kinds of observation statements.

Given these two principles, we have a way of arriving at observed error frequencies in the incorrigible metacorpus.

Given these observed error frequencies in a sample of observation reports, we can, other things being equal, infer long-run error frequencies among the various kinds of reports. Statements asserting that in general the frequency of erroneous raven judgments lies between (say) 0.0 and 0.1 may now appear in the metacorpus of evidential certainty.

What should be the basis of our choice between the two languages in question? One natural criterion would be the predictive observational content of our corpus of practical certainties $\mathbf{K}'$. That is, it seems reasonable to let past observations fend for themselves; theoretical understanding is notoriously difficult to distinguish from theological speculation; and prediction other than observational terms is hard to evaluate.

So what is "predictive observational content"? It is easy to find instances. Add to our metacorpus the sentence $\mathbf{O}(X$, "a_{17} is a raven", $t_{44})$, and see how many new sentences are added to our corpus of practical certainties. Generalize that idea: add to the incorrigible metacorpus a set of observation reports of each possible sort in equal numbers, and see what happens in the corpus of practical certainties. What happens depends on the values of p and p'; on what the original contents of the metacorpus is, on the language, and on the number of statements added. We will assume that this sensitivity to the levels of the evidential and practical corpora, and to the numbers of hypothetical observations added, is relatively unimportant or can be minimized. If this is so, we can adopt a third principle:

> $\mathbf{P}_3$ *The principle of maximizing predictive observational content:* Of two languages, adopt that that provides for the greatest predictive observational content in the corpus of practical certainties.

Applied to our example, this principle leads to the preferability of one language or another according to the relative frequencies with which various observations have been made. For example, if among our observation reports there have been a relatively large number of reports of nonblack ravens, then both "nonblack" and "raven" must be regarded as highly suspect or undependable if we were to speak $\mathbf{L}_2$. In particular, the situation might be so bad that, if one were to add $\mathbf{O}(X$, "a_{17} is a raven", $t_{44})$ to the incorrigible metacorpus, one could not even add "a_{17} is a raven" to the evidential corpus, and therefore neither that statement nor "a_{17} is black" to the corpus of practical certainties.

On the other hand, if almost all observed ravens have been reported to be black, the addition of the observation report to the metacorpus would lead to the addition of "a_{17} is black" to the practical corpus. Furthermore, the addition (in language $\mathbf{L}_2$) of $\mathbf{O}(X$, "a_{34} is not black", $t_{22})$ to the incorrigible metacorpus will lead (in $\mathbf{L}_2$) to the inclusion of "a_{34} is not a raven" in the corpus of practical certainties.

In the case in which very few nonblack ravens have been reported, $\mathbf{L}_1$ may lead to actions appropriate to the anticipation that an arbitrary new raven will be found to be black, on purely statistical grounds. (We will not in these circumstances be led to bet *against* that possibility at any odds.) But we will not be able to infer from the report that something is not black to the practical certainty that it is not a raven, and we will not be able to infer from the conjunction of reports that k objects are ravens to the conclusion that all k of them are black. Both of these inferences will go through if we adopt $\mathbf{L}_2$.

The upshot, for this simple example, is this: If practically all reports conform to the generalization that all ravens are black, $\mathbf{L}_2$ is the language of choice. If a significant number of reports conflict with that generalization, we must assume that our observations are so in error that $\mathbf{L}_2$ offers no advantages over $\mathbf{L}_1$, and indeed may suffer a loss in predictive observational content. (In fact, if the error rates are high enough, nothing may even get into the evidential corpus $\mathbf{K}_2$ on the basis of past observations.)

In summary, this approach generalizes in the following way:

fewer a priori constraints	more a priori constraints
observations without error little observational prediction no loss due to error	observations prone to error more observational prediction some loss due to error

3. Scientific Change

There are four cases of scientific change to consider. There is the addition of theory, there is the deletion of theory, there is replacement of one theory by another, and there is the addition to observational vocabulary. (The last mentioned is one that has not received much attention in the literature of scientific theory change, though it is obviously an important one.)

Let us first consider the addition of a theoretical generalization to a body of scientific knowledge. The theoretical generalization embodies new constraints. If these constraints lead to changes in our statistical knowledge concerning the errors of observation (or measurement), it can only be that we have been forced to allow for an increase in observational error.

Thus we assume the same observational vocabulary; we assume greater errors of observation; but in compensation, the predictive observational content of the corpus of practical certainties under the new a priori constraints will be larger than previously. We get new predictive content at a given level of practical certainty.

Note that there is nothing inductive about this. Our new generalization is not accepted because it has, relative to the evidence, a high degree of confirmation. In fact, its confirmation, relative to the body of knowledge expressed in the old language, is 0. (Relative to $\mathbf{L}_1$ the probability of "All ravens are black" is 0, as Carnap already recognized.[13]) Relative to the new language, its probability is 1.0: that is the probability of "All bachelors are unmarried," or "All ravens are black" in the language $\mathbf{L}_2$. Generalizations are not (directly) supported by their instances, but by their contribution to our foreknowledge. This may, but need not, be somewhat counterbalanced by a loss of precision in our predictive statements.

Note that the situation we face is the choice between two *given* languages. There is no question of using evidence to derive a new constraint. In this regard, also, then, the acquisition of new scientific knowledge is anti-inductive. We do not learn a new theory from experience; we test a new theory against an old one, by means of experience.

The second case to consider is that in which scientific progress consists in the deletion of an old generalization or law. I suspect that this is far more rare than some philosophers, Popper, for example[14], have maintained. Rather than refuting an old generalization, refined testing most often seems to lead to the replacement of one theory by another. But refutation no doubt does occur. When it does, what happens on the model I am depicting is that a group of new observation reports leads to new statistical error distributions. These new statistical error distributions may lead to a decrease in the observational content of the corpus of evidential certainties, and do lead to a decrease in the predictive observational content of the corpus of practical certainties. At this point, it may be that the language that *lacks* the generalization or law in question may lead to a practical corpus of greater predictive content.

Note that refutation does not occur as Popper and some other philosophers suggest. We do not put a theory to the test, note that the test results do not conform to the theory, and then reject the theory. Rather, an isolated test of the theory that does not yield the predicted result contributes to the data on the basis of which we derive our distribution of errors of observation. That this effect is not often pronounced (high school laboratory tests of the law of inertia do not contribute significantly to the refutation of Newtonian mechanics, nor to the statistics of error distribution in the measurement of mass and distance) does not show that it does not exist. In some cases there is an extremely large body of data on which the error distribution is based. In such cases it can *look* as though a theory is being refuted. Note also that according to the conventional treatment of error in physics, *no* set of measurement results is ruled out as impossible by any theory. If we take error into account, no experiment can refute a theory.

The third and most common form of theoretical change is the replacement of one theory by another. Again, we assume that the observational vocabulary is the same in each case. We compare directly the error statistics and the corresponding predictive observational contents of the corpus of practical certainties of the two theories.

Only one philosopher has made a serious attempt to account in rational terms for the replacement of one theory by another. This is Isaac Levi.[15] For Levi the problem is particularly difficult, since on his view (as on mine), within the framework provided by one theoretical language, direct refutation is impossible. If an experiment is performed, and the results contradict the predictions of the theory, it is the results that must be rejected as false, not the theory. On my view, the erroneous results contribute statistically to our knowledge of the theory of error of the kinds of observations involved.

Levi's solution is that when things get bad enough, we may be led to contract to a body of knowledge in which neither the original theory nor its rival appears. When we weigh Newtonian mechanics against relativity mechanics, we first contract to a body of knowledge with no mechanics. Then we use our evidence as a basis for new expansion. Replacement is thus construed as contraction, followed by expansion.

On the view that I am advocating, this two-stage process is not necessary. We may compare directly the predictive observational contents of the corpora of practical certainties of the two cases, corresponding to the two languages.

The fourth and final case to consider is when the two languages differ in the observational vocabulary. Since I have rejected the incorrigibility of observation statements, it is necessary to say a bit about what a change in observational vocabulary comes to. What constitutes an observational term is dependent on the corresponding error theory and on what constitutes the level of evidential certainty. The basic idea is that a term is observational if an observation report making use of that term constitutes adequate evidence for including the corresponding statement among the evidential certainties of a given level of acceptance.

The kind of change in vocabulary I have in mind here includes the South Seas islander learning to distinguish dependably among the kinds of snow, or the Eskimo learning to distinguish dependably among the varieties of wave patterns. An even simpler and more direct example, however, is provided by scientists learning to see through a microscope or telescope. Galileo's inquisitors, after all, didn't need to refuse to look through his telescope. They could have looked through it, and what they would have seen were flashes of light, their own eyelashes, and difficult-to-identify shapes. It takes the student of histology a significant period of time to learn to see through a microscope. Observation, like any other skill, must be learned.

The reward of learning to observe, whether we consider the individual or the group of individuals, is the same as always: an increase in the predictive observational content of the corpus of practical certainties.

If we look at the change of observational vocabulary as learning to see something new, we may suppose that the old observational predicates are included in the new ones; what we achieve is a *refinement* of our observational vocabulary.

This case is somewhat more complicated than the preceding three cases, since we may consider two kinds of comparisons between the old and the new corpus of practical certainties. We may consider the predictive observational content of the practical corpora, constrained to sentences in the old vocabulary. Or we may look simply at the absolute numbers of predictive observational consequences in the two cases. The latter approach is not as question begging as it might appear to be, due to the requirement that *observational* predicates must be predicates that we can dependably apply. Thus one may imagine (what is no doubt the case) that phrenologists could learn to distinguish head bumps that modern people can no longer dependably detect. Well and good. But the predictive observational content of the phrenologist, given the addition of bump reports in the metacorpus of incorrigibilia, would presumably not exceed that of his modern nonphrenological counterpart, even if we include bump reports among the predictive observation statements to be taken account of.

4. Conclusions

This view of scientific knowledge and scientific change bears on a number of contemporary issues.

It bears on the question of incommensurability of alternative theories, for example. Kuhn,[16] Feyerabend,[17] and others have sometimes written as if there were no way of directly comparing theories and their successors. If the view of theories as mere conventions that I have been outlining is meritorious, it is easy to see the truth in such a view. It is quite true that two competing theories are not comparable in the sense that we can look at the evidence and decide that one of them has (or should have) a larger degree of confirmation than the other, or that one has been tested and refuted while the other has not, or that one has been tested more severely, or that one is more testable than the other. Two theories are not compared directly in the light of the evidence. But given a body of observations—that is, given a set of **O**-statements in the incorrigible metacorpus—we can find a direct basis of comparison of two *languages* in the predictive observational content of the corpus of practical certainties.

A serious question that might be raised concerns the apparent ubiquity of language change in the history of science. As I have sketched scientific change, all change, other than mere statistical inference, reflects a change in the language of science. Is it really the case that every time a universal generalization is confirmed (as we naively say), what is going on is that we are choosing between languages?

We must distinguish two cases. The natural history case, in which we look at ravens or swans and arrive at a universal generalization, it seems to me, does involve change of language. It is not that we decide that "almost all swans are white," and then discover that Australian swans are black. It is that we take as a matter of certainty (convention) that all swans are white; and then, given the evidence from Australia, must decide – nontrivially – whether there are adequate anatomical grounds for regarding the southern black birds as also swans. It is quite clear that there are issues in such a case as this that only an ornithologist is in a position to deal with.

The other case is the one in which we already have a body of structured scientific knowledge to use. This corresponds to what used to be called "demonstrative induction." We can decide that all samples of the compound X melt at about 45 degrees Celsius, on the basis of a small number of experiments, because it is a part of the general theory—i.e., language—that each pure compound has exactly one melting point. All we have to do is to determine that one sample of X melts at 45 degrees to determine that all samples do.

Can we decide between *any* two languages for science? For example, between

French physics and English physics? The answer is clearly negative. French physics and English physics (assuming they embody the same laws) will have the same number of predictive observational consequences in the corpus of practical certainties. Indeed, they will have the same predictive observational consequences, though expressed in a different vocabulary. So the difference between French physics and English physics is exactly a matter of *arbitrary* convention. The difference between Newtonian physics and Einsteinian physics is also a matter of convention, but while it is *mere* convention, it is not arbitrary convention. Einstein's physics allows us to predict a bit more.

This way of looking at scientific languages might be thought to have consequences for the issues of scientific realism and instrumentalism. This is true, but it might also be thought that since I am arguing that scientific theories are mere conventions, instrumentalism is the only plausible ontological view. This is mistaken. It reflects a confusion between mere and arbitrary convention. If the difference between theories quantifying over different sorts of entities were a matter of arbitrary convention, then perhaps instrumentalism would be warranted. But I have just argued that the choice between alternative scientific languages is an epistemological issue. We have epistemological grounds for choosing one theory over another. And this suggests that we can have epistemological grounds for quantifying over one set of entities rather than another.

In addition, the framework suggested weakens the distinction between "observational" and "theoretical" terms profoundly. Given modern instrumentation and modern training, it seems clear that many of the properties of cells must be regarded as observational properties. Before the invention of the microscope, some of these same properties must have been taken to be unobservable and theoretical. There are indeed complexities to be dealt with here, but the point that emerges quite clearly is that nothing definitive—in particular no form of instrumentalism—emerges as entailed by the view described.

To view scientific theories as mere but not arbitrary conventions offers a framework in which the rational succession of one scientific theory by another can be understood; it offers an alternative to the view of scientific theories as subjective, arbitrary, and subject to personal whim and social pressure, rather than the demands of rationality. What views of the world the evidence we happen to have supports is an objective, rational matter. This is not to say that experience and reason dictate what theory we should hold of the world, but that experience and reason dictate which of two *alternative* theories (languages) is preferable.

Notes

1. Ernst Mach, *The Science of Mechanics* (LaSalle, Ill.: Open Court, 1960). 1st German ed., 1883; 1st American ed., 1893.

2. Henri Poincaré, *Science and Hypothesis* (New York: Dover Publications, 1952). 1st French ed., *La Science et L'hypothese* (Paris, 1903).

3. Pierre Duhem, *The Aim and Structure of Physical Theory* (Princeton: Princeton University Press, 1954). 1st French ed., 1906.

4. W. V. O. Quine, "Two Dogmas of Empiricism," *The Philosophical Review* 60 (1951). This essay has been extensively reprinted. It appears in Sandra Harding ed., *Can Theories Be Refuted?* (Dordrecht: Reidel, 1976), 41–64. This volume contains a number of interesting essays that bear on our present concerns.

5. W. V. O. Quine and J. S. Ullian, *The Web of Belief* (New York: Random House, 1970). This is not the first appearance of Quine's web metaphor, but this volume contains the most developed statement of the metaphor.

6. William Craig, "Replacement of Auxiliary Expressions," *Philosophical Review* 65, 1956, 38–55.

7. W. V. O. Quine, *Word and Object* (Cambridge: MIT Press, 1960). The issue is related to the question of radical translation; can we sort our experiences uniquely according to the expressions of our language?

8. Quine's *Word and Object* is the locus classicus for underdetermination arguments. It should be noted that Quine's own views have evolved significantly since that work.

9. David Lewis, *Convention: A Philosophical Study* (Cambridge: Harvard University Press, 1969).

10. Henry E. Kyburg, Jr., "The Reference Class," *Philosophy of Science* 50 (1983): 374–97. The issue has been addressed in a number of other publications, including my *The Logical Foundations of Statistical Inference*. (Dordrecht: Reidel, 1974).

11. Henry E. Kyburg, Jr., "Full Belief," *Theory and Decision* 25, 1988, 137–162, contains one proposal for the determination of p and p'. Others are no doubt possible.

12. Mary Hesse, *The Structure of Scientific Inference* (Berkeley and Los Angeles: University of California Press, 1974).

13. Rudolf Carnap, "On Inductive Logic," *Philosophy of Science* 12, 1945, 72–97.

14. K. R. Popper, *The Logic of Scientific Discovery* (London: Hutchinson and Co., 1959). The first German edition appeared in 1934, and Popper has had a numerous and influential following ever since.

15. Isaac Levi, *The Enterprise of Knowledge* (Cambridge: MIT Press, 1980). Recent work on this problem has also been done by Peter Gardenfors and others; see *Knowledge in Flux* (Cambridge, Mass.: Bradford Books, 1988) and references therein.

16. Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

17. Paul K. Feyerabend, "Against Method: Outline of an Anarchistic Theory of Knowledge," in *Minnesota Studies on the Philosophy of Science*, Vol. 4, *Analyses of Theories and Methods of Physics and Psychology*, eds. Michael Radner and Stephen Winokur (Minneapolis: University of Minnesota Press, 1970), 3–130.