**Rational Belief and Probability Kinematics**

Bas C. Van Fraassen

*Philosophy of Science*, Vol. 47, No. 2 (Jun., 1980), 165-187.

Stable URL:
http://links.jstor.org/sici?sici=0031-8248%28198006%2947%3A2%3C165%3ARBAPK%3E2.0.CO%3B2-2

*Philosophy of Science* is currently published by The University of Chicago Press.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Sun Mar 6 03:07:14 2005

# RATIONAL BELIEF AND PROBABILITY KINEMATICS*

BAS C. VAN FRAASSEN†

*University of Toronto and University of Southern California*

A general form is proposed for epistemological theories, the relevant factors being: the family of epistemic judgments, the epistemic state, the epistemic commitment (governing change of state), and the family of possible epistemic inputs (deliverances of experience). First a simple theory is examined in which the states are probability functions, and the subject of probability kinematics introduced by Richard Jeffrey is explored. Then a second theory is examined in which the state has as constituents a body of information (rational corpus) and a recipe that determines the accepted epistemic judgments on the basis of this corpus. Through an examination of several approaches to the statistical syllogism, a relation is again established with Jeffrey's generalized conditionalization.

In this paper I shall describe what I take to be the general form of an epistemological theory and consider two simple (but mutually very different) sorts of theories of this form. In both cases, adaptations of Richard Jeffrey's probability kinematics play a central role.

**1. Epistemological Theories.** By an *epistemological theory* I mean a theory which provides a representation of the main epistemic factors; and these I take to be as follows. There is a family of *epistemic judgments* which comprise all expressions of full, partial and comparative belief or assent. Each person has at any given time an *epistemic state,* functionally related to the epistemic judgments which he makes, or rather to which he is committed at that time. He has also an *epistemic commitment,* a function determining how, at least as presently intended, the epistemic state would change under certain conditions. The latter conditions, the deliverances of experience so to say, form the *epistemic input.* Each of these four items is an epistemic factor. In addition to their representation, the theory must also advance *criteria*

*Received June 1979; revised October 1979.

†Research for this paper was supported by NSF grant SOC78-08464. I also acknowledge gratefully correspondence and discussions with Dorling, Giere, Harper, Jeffrey, Kyburg, Levi, Lewis, Salmon, Seidenfeld, Williams, and Zanotti, and the helpful comments made by Domotor on the penultimate draft. Since the correct representation of the views of some authors turned out to be a delicate matter, I emphasize that even when I associate ideas with their names, the connection was suggestive, and I am solely responsible for the shortcomings.

*Philosophy of Science,* 47 (1980) pp. 165–187.
Copyright © 1980 by the Philosophy of Science Association.

*of rationality* by which the state and commitment are evaluated.

Being prior to the development of theory, the above remarks do not constitute definitions, but are meant as an intuitive introduction of the main concepts. I shall explore some kinds of epistemological theory, admittedly simple and only partially developed. In this way, I hope to throw light on the general structure of epistemological theory and on the possibilities open to it.

**2. Epistemic Judgments.** As expressed in language, our beliefs reveal a bewildering variety. A person may say that he believes that $A$, or that he is fairly convinced that $A$, that he would not bet that $A$, that it seems more likely to him that $A$ than that $B$, or that, assuming $C$, it seems as likely as not that $A$. This series of schematic examples is already rather far removed from the data, and it is not at all clear that it does justice to those data. A theory will provide, explicitly or implicitly, a typology of such judgments, and aim to make sense of the ones that fit its catalogue.

The simple, deliberately restricted, typology which I shall use in this paper is as follows. Epistemic judgments are made up from an underlying set $F$ of propositions, which include a tautology $T$. There are four types of epistemic judgments; their English readings are "It seems less likely that $A$ than that $B$"; "It seems no less likely that $A$ than that $B$"; "The probability (to me) that $A$ is no less than $r$, no greater than $s$"; "The probability (to me) that $A$ is not between $r$ and $s$ (inclusive)." So the second and fourth types are the contraries of the first and third. An expression of full belief I shall equate here with a judgment of second type in which $B$ is a tautology.

Merely giving the English readings does not yield a typology; for those readings are at most clues to the intended relations among those judgments. To represent those relations we must use a set of *valuations* of the underlying propositions. These valuations must at least have the property that they assign each proposition in $F$ an element of some partially ordered structure, *and* a numerical interval. (The two may be identical, since those intervals are partially ordered in the requisite way; and numbers we may regard as degenerate intervals. Therefore, the valuations *could be* probability functions defined on $F$; that is the most obvious example of valuations.) In addition those valuations may be required to respect the structure of the propositions in certain respects. Hence part of the specification of the typology is the selection of a certain set of valuations, the *admissible epistemic valuations*. Relations among epistemic judgments derive from these in obvious ways. Such a valuation *satisfies* "it seems less likely that $A$ than that $B$" exactly if it assigns $A$ as first element something that precedes (in the partial ordering in question) the first element it assigns to $B$; it *satisfies* "the probability (to me) that $A$ is no less than $r$, no greater than $s$" exactly if the numerical interval it assigns $A$ is part of $[r,s]$. The contrary of a judgment is satisfied exactly if the judgment itself is not. One epistemic judgment *implies* another exactly if all the admissible epistemic valuations that satisfy the first also satisfy the second; and so forth.

We may expect the theory to envisage some relationship between valuations and epistemic states, and perhaps aiso between valuations and truth conditions, frequencies, or chance. But *prima facie* such relationships need not be very close. The valuations determine the logical relationships among the epistemic judgments, and by varying the set of valuations these relationships can be made as strict or as loose as we see fit. The typology I have given is simple, but the method of setting it up generalizes readily to more extensive classifications.

**3. Epistemic States: The First Theory.** The simplest reasonable theory that comes to mind in view of the preceding (and the literature) is this:

(a) the set $F$ of propositions is (representable as) a field of subsets of a given set $T$,
(b) the admissible epistemic valuations are the probability functions defined on $F$,
(c) the epistemic states are the admissible epistemic valuations.

The person is committed to an epistemic judgment at a given time if and only if his epistemic state at that time satisfies the judgment. It will be noticed therefore that his state at a time, and the judgments to which he is committed at that time, can each be determined from the other. There are no "hidden variables" in this theory: the state can be determined uniquely if we can determine the judgments. To complete the theory, we must describe the possible epistemic commitments, which govern the change of state in response to epistemic input. The first item on the agenda is therefore the representation of such input.

**4. Epistemic Input: A First Typology.** How shall we describe the deliverances of experience? If we were concerned not with epistemology generally but with statistical practice, we could simply say: some proposition $E$ is received as new evidence.[1] For a statistician, *qua* statistician, no other aspect of experience would be relevant. Nor would there be any question as to what that evidence $E$ is; he is employed to accept certain data as input for his calculations. If we extrapolate from this paradigm to epistemology generally, we get what I shall call the *revelation model* of belief change.

The epistemic commitment generally advanced as rational for such input is given by the familiar formula

$$\text{(C)}\quad P'(A) = P(A|E)$$

which yields the posterior probability $P'$ for a prior probability $P$ which assigns a positive value to $E$, when $E$ is the total evidence. I shall call this shift from $P$ to $P'$ *simple conditionalization;* that it is the correct formula, Hacking originally called the "Bayesian dynamic assumption" (1967). (Note: I shall write "$P(A|E)$" or "$P_E(A)$" for "$P(A \& E)/P(E)$" and use without comment will imply that $P(E)$ is positive.)

This representation of the deliverance of experience as a proposition in $F$ is an extreme position. A contrary extreme position denies that the input can be represented in terms of propositions at all. I take this to be intended when the input is described as "sensation," or an analogy is drawn with a robot capable of learning. The deliverance of experience, it is then said, is pre-conscious, or at least pre-judgmental. There is no "reasoning oneself into" a new epistemic state, since the state would already have changed when the propositions needed as premises had been accepted, or when the epistemic judgments needed as premises, had been arrived at in some other way. The epistemic commitment then does not operate on the conscious level; it is the function that takes sensation and epistemic state into new epistemic state, and that this has happened shows up at the conscious level in changed epistemic judgments. Sensation, epistemic input, is then to be represented as a quantitative variable. Let us call this the *learning machine model* of belief change.

There is a major difference between the two approaches, since in the second, the question of justification effectively disappears. As normally used, the question "why did you change your mind?" has two senses. It could be a request either for explanation or for justification. Within the revelation model, both senses make sense. Asked why he now thinks it likely that $A$, the person can say "because

[1]When two spaces are used (parameter or hypothesis space and observation or sample space) I shall think of these as subspaces of a larger one (possibly produced by a product construction), so that in a single context all propositions are represented by measurable sets in a single space.

I received evidence $E$." We, and he, can then investigate whether he has responded rationally to the total evidence he received; specifically, did he follow formula (C)? On the learning machine model, the justification sense of the question disappears, since any judgment to the effect that $E$ is the case, is part of the response to the input, and that one's sensation had this character or that would also be a judgment which is part of that response. The criteria of coherence among one's present judgments would apply, of course, but no criteria of rationality concerning the changes in response to the deliverance of his experience are available to the person.

As third alternative I propose that the inputs, though not representable *as* propositions, are representable *in terms of* propositions. The idea is that in response to what happens to him, the person accepts not propositions, but *constraints* on his posterior epistemic state. He can phrase these as *commands* to himself (the simplest being "have full belief that $E$!"). These commands are the epistemic input, and his epistemic commitment maps such input, plus prior epistemic state, into posterior state.

In this way some questions of justification arise. Asked why he now thinks it likely that $A$, he can reply: in response to my experience, I accepted constraint **C** on what my epistemic judgments should be, and then I satisfied that constraint by changing my epistemic state in manner $Y$, and part of the result was this new judgment concerning $A$. We can then investigate his rationality in two ways: did the shift in manner $Y$ indeed satisfy constraint **C**, *and* was that a rationally optimal way of satisfying the constraint? What we still cannot do, of course, is ask him to justify his acceptance of that constraint in response to the experience he had—but similarly, in the revelation model there is no justification for why $E$ was accepted. (Though experience speaks with the voice of an angel.)

Since there are many sorts of constraints that can be imposed on posterior epistemic states, we require also a typology of inputs. Still following the simple theory, I propose as simple typology a classification which I can introduce by means of some examples. In the present simple theory, the set of propositions $F$ is (represented as) a field of subsets of a given set $T$ and $T$ can be taken to be the sample space for the probability functions. Hence the following all make sense:

(i) the posterior probability of $E$ should be 1;
(ii) the posterior probability of $A$ should be $q(A)$ for each member $A$ of set $X$;
(iii) the posterior conditional probability of $A$ given $B$ should be

$q(A,B)$ for each couple $(A,B)$ in family $Y$ of such couples;

(iv) the posterior expectation of $g$ should equal $E(g)$ for each member $g$ of family $Z$ of random variables.

The admissible valuations in the representation of epistemic judgments will here play the role of demarcation of *possible* inputs. In the present case that means, for instance, that a command of type (i) does not represent a possible input if $E$ is a self-contradiction.

The four types are not disjoint but of increasing generality. The first clearly corresponds to the sole form of input that occurs in the revelation model. The second corresponds to the sorts of input discussed by Richard Jeffrey (1965) in connection with his generalized conditionalization formula (of which (C) is a special case):

$$\text{(J) } P'(A) = \Sigma\{q(B)P(A \mid B)\text{: } B \text{ in } X\}$$

applicable only if $X$ is a *partition* of the sample space, each of whose members has a non-zero prior probability. The third generalizes that in an obvious fashion; but (i)-(iii) are all special cases of (iv). For the conditional probability $P'(B \mid C)$ will be $q(B,C)$ exactly if that is the expectation value of the variable which takes value $q(B,C)$ on the complement of $C$, value 1 on the intersection of $B$ and $C$, and value zero elsewhere.

We have now arrived at a reasonable generalized formulation of the subject of *probability kinematics,* which Jeffrey introduced. The next question we face is: how is the posterior probability $P'$ determined by the prior probability $P$ and the constraint that represents the epistemic input at that time? That is the question: *what is the epistemic commitment function?* Since we are engaged in epistemology rather than psychology, this is focused to: *what is a rational epistemic commitment function?* What forms does the change from prior to posterior epistemic state take, and how is it justified or justifiable in specific cases? I shall broach these questions in the first place through an exploration of the kinds of change that satisfy equation (J) above.[2]

**5. Probability Kinematics.** Let us consider how two probability functions $P$ and $P'$, defined on the same field, may be related to each other. Call $P'$ *absolutely continuous* with respect to $P$ exactly if $P'$ assigns zero to all propositions to which $P$ assigns zero. I shall consider no other cases in this section ("zeros are not raised" in changes

[2]In view of recent controversies (e.g. Kyburg 1977, Levi 1977) I should emphasize that there are several principles which may reasonably be called "conditionalization" and that acceptance of the present typology of inputs implicitly rejects some of these.

to posterior probability $P'$ to be discussed here). If the field is finite, the following will necessarily hold then:

(JS) There is a finite measurable partition $X$ and non-negative numbers $q(A)$ for $A$ in $X$ summing to 1, such that $P'(B) = \Sigma\{q(A)P(B/A): A \text{ in } X\}$.

For if the field is finite, and $P'$ absolutely continuous with respect to $P$, then such a partition $X$ can be formed by taking all but one atom of the field to which $P$ gives a positive value, and the union of all remaining atoms.

If the field is not finite we have no such guarantee. Let us call the transformation of $P$ into $P'$ described in (JS) a *Jeffrey shift* with *base* $X$. (Note: it will be part of the meaning of "Jeffrey shift" here that the base is finite.) Let us call such a shift *pure* if none of the coefficients $q(A)$ is zero, and call it *prime* if all but one of them is zero. A prime shift is the same as what I earlier called a simple conditionalization. We may note about these shifts that (a) they are partial operations on the family of probability functions defined on the field (since it is required that $P$ give a positive value to each member of the base); (b) each is equal to a weighted sum of prime shifts; and (c) each Jeffrey shift is equal to a pure shift followed by a prime shift. These remarks go some way toward justifying separate discussion of pure shifts and of prime shifts.

**5.1 The Identification Problem.** When an epistemic state changes from $P$ to $P'$, which Jeffrey shift (if any) was applied? There are usually many that lead from $P$ to $P'$, if any one does; since the change can be redescribed using any finer partition as base. I came to think of this as a problem after reading Levi (1967, Section III). For if the person is asked for a justification ("Why did you change to $P'$ from $P$?") he must surely first identify what change he made. The answer at which I arrived is that there is always a "minimum" Jeffrey shift that can be identified as having effected the change, if any has.

Let us call partition $X'$ a *refinement* of (or *at least as fine as*) partition $X$ (and the latter a *coarsening* of, or *at least as coarse as* the former) exactly if every member of $X'$ is a subset of some member of $X$. When $X$ is the base of Jeffrey shift $f$, and $P' = f(P)$, call $X$ also a *P-base* of $P'$. If the Jeffrey shift is pure (and in this subsection I shall consider only pure cases), this is symmetric in $P$ and $P'$. Given $P$, $P'$ and $P$-base $X$ of $P'$, the shift $f$ with base $X$ that leads from $P$ to $P'$ is uniquely determined. What I mean to show is that if $P'$ has any $P$-base, it has a coarsest one.

**5.2 The Coarsest *P*-base.** Conditionalization, whether ordinary or Jeffrey, is essentially orthogonal decomposition. Call $P$ and $P'$ *orthogonal* if they have the same domain and $P(A) = 1$ while $P'(A) = 0$, for some $A$. The equation

$$P = \Sigma\{a_i P_i : i \text{ in } I\}$$

is an *orthogonal decomposition* of $P$ exactly if $I$ is countable, the numbers $a_i$ are non-negative and sum to 1, and the probability functions $P_i$ are all mutually orthogonal. Such a decomposition exists exactly if there exists a countable partition $X$ such that $X = \{A_i : i \text{ in } I\}$ and

$$a_i = P(A_i);\ P_i = P(-|A_i) \text{ when } a_i \neq 0$$

for all $i$ in $I$ (See van Fraassen 1979, Section 5.) Thus a Jeffrey shift is a special case of orthogonal decomposition of the posterior probability: all the $P_i$ equal the prior probability. At the same time therefore, that prior can be decomposed on partition $X$ as well.

Call $A$ a *P-eigenproposition* of $P'$ exactly if $P(-|A) = P'(-|A)$, both being well-defined; then partition $X$ is a $P$-base for $P'$ exactly if $X$ consists of $P$-eigenpropositions of $P'$. Note that the relation is symmetric. Finally, call $A$ and $B$ *compatible for P and P'* exactly if their union is a $P$-eigenproposition of $P'$. The following are equivalent when $A$ and $B$ are disjoint $P$-eigenpropositions of $P'$:

1. $A$ and $B$ are compatible
2. $P(A|A \cup B) = P'(A|A \cup B)$ and $P(B|A \cup B) = P'(B|A \cup B)$
3. $P(A)/P'(A) = P(B)/P'(B)$
4. The $P$-odds of $A$ to $B$ equal the $P'$-odds of $A$ to $B$, i.e. $P(A)/P(B) = P'(A)/P'(B)$

The last condition is perhaps the most intuitive: for members of a $P$-base, *compatibility is the same as equality of prior and posterior odds.* In (2), as in all similar assertions, it is to be understood as implied that the terms are well-defined.

The equivalence of (3) and (4) is trivial, and that of (2) and (3) hardly less so, given that $A$ and $B$ are disjoint. To prove the equivalence of (1) and (2) we note that $A$ and $B$ are disjoint and use the general principle $P(E \cap F|E \cup D) = P(E|E \cup D)\, P(F|E)$:

$$P(X|A \cup B) = P(A|A \cup B)P(X|A) + P(B|A \cup B)\,P(X|B)$$

$$P'(X|A \cup B) = P'(A|A \cup B)P'(X|A) + P'(B|A \cup B)\,P'(X|B)$$

Because $A$ and $B$ are $P$-eigenpropositions of $P'$, the lefthand quantities are equal if and only if (2) holds.

From this first lemma we draw the consequences that compatibility is transitive, and also that if $A_1, \ldots, A_n$ are disjoint $P$-eigenpropositions of $P'$, while $A_i$ is compatible with $A_{i+1}$, for $i = 1, \ldots, n - 1$, then the union of these propositions is also a $P$-eigenproposition of $P'$. Thus compatible members of a $P$-base can be summed arbitrarily to produce another $P$-base. (If $P$, $P'$ are countably additive, this remark generalizes to the countable case.)

The second lemma is that if $X$ and $X'$ are finite $P$-bases for $P'$, they have a common coarsening which is also such a base. For let $X$, $X'$ be thus. They have a coarsest common refinement $Y$ which is still obviously a $P$-base, namely

$$Y = \{A \cap B : A \text{ is in } X, B \text{ is in } Y, A \text{ and } B \text{ are not disjoint}\}.$$

We now coarsen $Y$ by unionizing any and all compatible elements of it. By the preceding lemma this can be done in arbitrary order, and produces still another $P$-base. Since $Y$ too is finite, this process comes to a stop; call the result $X''$. But if $A$ is in $X$ then $A$ is the union of $\{A \cap B : B \text{ in } X'\}$, whose elements are compatible, and which appears therefore at least as a subset of some element of $X''$. Therefore *$X''$ is a coarsening of $X$*; and by similar reasoning, of $X'$ as well.

This leads then to the theorem: *If $P'$ has any finite $P$-base then it has a coarsest one.*[3] For let $Q$ be the set of all finite $P$-bases of $P'$ and assume it non-empty. $Q$ is partially ordered by the refinement relation. Consider any chain $R$ in (linearly ordered subset of) $Q$. The cardinalities of its members form a strictly decreasing sequence of natural numbers, and must therefore have a last member. So every chain in $Q$ has an upper bound in $Q$ (its own last member), and by Zorn's lemma, $Q$ has maximal (coarsest) elements. If $X$, $X'$ are two such maximal elements, however, they have a common coarsening by the second lemma—so there cannot be more than one such maximal element.

**5.3 The Justification Problem.** Suppose the agent has shifted from prior $P$ to posterior $P'$, and has identified this as an instance of Jeffrey shift $f$ (whose base is the coarsest $P$-base of $P'$). We ask him now to justify this move, or at least to explain why he thinks that this is what he should have done, or failing that, why he considers

[3]When these notes were discussed in Jeffrey's seminar, Persi Diakonis pointed out that this theorem follows also from general results concerning the existence of minimal sufficient statistics. The relation between bases and sufficiency is easily seen (as Zoltan Domotor later remarked) from the definition of the partial pairwise sufficiency of partition $X$ for $P$ and $P'$ by the equation $\Sigma\{I_A(P_A - P'_A): A \in X\} = 0$.

this shift to be at least as good as any others he could have made, and so forth.

The general form of his answer should surely be this, at least, given the typology of inputs we have available: he claims that in the situation, in view of his experience, he accepted certain constraints on his posterior epistemic state, that his shift gave him a posterior that does satisfy those, and that he managed to satisfy them in a manner that was optimal in some or certain respects. Indeed, the optimality will lie in his assertion that he did not jump to conclusions, or renounce earlier epistemic judgments capriciously: that the shift constitutes a minimal change needed to satisfy the constraints.

It was apparently first proved by B. Jamison that Jeffrey's rule (J) minimizes relative information in $P'$ with respect to $P$ subject to the constraint type (ii) in section 4 above.[4] This was reported by May and Harper (1976, pp. 141–143) before publication; (see Jamison 1974). Since then this point was developed by P. Williams (1978) and by Domotor, Zanotti, and Hagen (1979); all note that it follows from general results in information theory, but the last two papers use it to provide interesting and general clarifications of probability kinematics. I shall give a brief exposition (which I circulated in notes before seeing the last two papers).

Using essentially Jaynes' definition (see Hobson 1971, chapter 2) the information in $P'$ relative to $P$, as measured in partition $X$, equals

$$I(P',P,X) = \Sigma\{P'(A) \log (P'(A)/P(A)) \colon A \text{ in } X\}.$$

This is always non-negative, zero only if $P$ is the same as $P'$, and satisfies the decomposition principle:

If $X'$ is a refinement of $X$ then $I(P',P,X') =$

$I(P',P,X) + I(P', P, X|X')$,

where the conditionalized quantity $I(P', P, X|X')$ equals $\Sigma\{P'(A)I(P'_A,P_A, X')\colon A \text{ in } X\}$. This is also always non-negative. It follows at once from this principle that if $X'$ is a refinement of $X$, then the relative information as measured in $X'$ is at least as great as it is measured in $X$.

Suppose that $P'$ must satisfy the constraint that $P'(A) = q(A)$ for each $A$ in partition $X$. Then all posterior $P'$ satisfying the constraint, have the same amount of information relative to $P$ as measured in

[4]Domotor has suggested that the credit should perhaps go to Jaynes, on the basis of the opening article of R. D. Levine and M. Tribus (eds.) *Maximum Entropy Formalism Conference* (Cambridge: MIT Press, 1979). My own acquaintance with the idea came from May and Harper (1976).

$X$. As measured in finer partitions, the amount of relative information may differ. But we see at once that if $P_A = P'_A$ for each $A$ in $X$, then the decomposition principle implies that $I(P', P, X) = I(P', P, X')$ for any refinement $X'$ of $X$. Hence in that case, the relative information is at a minimum as measured in any refinement of $X$—the case in which equation (J) holds for exactly the partition and number involved in the constraint.

**5.4 Generalization of Constraints.** As pointed out in Section 4, constraints on posterior probability can take the more general forms. Jeffrey himself mentioned the case in which $P'(B) = q(B)$ for each member $B$ of some family $Y$, which is not a partition. Any such case can be put in the form of constraint type (iv), in which random variables are required to have given posterior expectation values. I shall not discuss this in full generality; restricting myself to finite families of simple random variables.

A *simple* random variable is a finite linear combination of characteristic functions, $g(x) = \Sigma\{a(A)I_A(x)$: $A$ in finite partition $X\}$. Let us begin with a simple such variable, and assume that $P(A)$ is positive for all $A$ in $X$, and that $X$ has more than one member. The constraint is that

$$\Sigma\{P'(A)a(A): A \text{ in } X\}$$

equals a given number $r$. We wish to find the probability $P'$ satisfying this constraint that has minimum information relative to $P$, as measured in any refinement of partition $X$. Step one is clearly to make it a Jeffrey shift, making $X$ a $P$-base of $P'$, thus reducing the problem to the previous case. It remains then to find the posterior probabilities for members of $X$.

Let the prior probabilities of these propositions be $p_1, \ldots, p_n$ and the posterior be called $x_1, \ldots, x_n$, and let the corresponding values $a(A)$, for $A$ in $X$ be $k_1, \ldots, k_n$. The variables $x_i$ range continuously over the interval $[0,1]$. The functions

$$f(x_1, \ldots, x_n) = \Sigma x_i \log (x_i/p_i)$$

$$g'(x_1, \ldots, x_n) = (\Sigma x_i k_i) - r$$

corresponding to relative information, and to the constant zero respectively (if the constraint is satisfied), are continuously differentiable on the open interval $(0,1)$, so the Lagrange Multiplier Theorem applies. That is, the extrema of $f$ are found at the points where $grad\ f = m\ grad\ g'$, $m$ being a constant.

$$grad\ f = \left\langle \frac{\partial}{\partial x_1} f, \ldots, \frac{\partial}{\partial x_n} f \right\rangle$$
$$= \langle \ldots, \log x_i + 1 - \log p_i, \ldots \rangle$$
$$grad\ g' = \langle \ldots, k_i, \ldots \rangle$$

so we can solve $m\ grad\ g' = grad\ f$ to get that

$$x_i = e^{mk_i} p_i / e$$

So far, however, I have ignored the obvious further constraint that the posterior probabilities in $X$ must sum to 1. Hence all the $x_i$ must be divided by their sum:

$$q_i = x_i / \Sigma x_i = e^{mk_i} p_i / Z$$

where $Z$ is the normalization factor. In any concrete case, $m$ can be determined numerically, by considering the possible ways in which $grad\ f$ can be a multiple of $grad\ g'$.

The solution to our problem was therefore this: to set the posterior expectation value of $g$ equal to $r$, while minimizing relative information, utilize the Jeffrey shift on base $X$ (used to define $g$), with the coefficients as determined above.

Generalizing to a finite family $Y$ of variables uses the Lagrange Theorem with constants $m_i, \ldots, m_s$ and the solution

$$q_i = e^{\Sigma m_j k_i^j p_i} / Z'$$

where $k_i^j$ is the value taken by the $j^{th}$ variable in $Y$ on the $i^{th}$ proposition in partition $X$, and the sum is over index $j = 1, \ldots, s$.

We have seen therefore that in a broad range of constraints (much broader than those initially discussed by Jeffrey), it is possible to choose a (pure) Jeffrey shift that yields a posterior satisfying that constraint while minimizing the information relative to the prior. (Note, however, that I specifically excluded from discussion those cases in which simple conditionalization (C) would occur; not because I wish to rule them out but because Jeffrey shifts in general can be regarded as composed of prime and pure shifts which can be made in succession.)

**5.5 Field's Re-Parametrization.** In a recent study (1978) Hartry Field suggests that the deliverances of experience, since they lead to a transformation of prior into posterior probability, should themselves be characterized by a parameter that determines that transformation. In the view I have developed so far, that shift is determined by two factors: the deliverance of experience (represented by a command imposing a constraint on the posterior) plus the epistemic commitment (which consists, in the simple theory, merely of a wish to make the change optimal in some respect, e.g. minimize relative information). Field's parameter represents therefore, in my view, the combined effect of two factors.

The epistemic input as conceived by Field specifies both a partition $X$ and a parameter $b$ taking real values on that partition which sum to zero. Only pure Jeffrey shifts are treated directly again; the partition $X$ is therefore not the trivial (one-member) one. Field's formula for the posterior probability of $A$ in $X$ is

$$\text{(F) } P'(A) = P(A)e^{b(A)}/Z$$

where denominator $Z$ is the normalization factor (sum of all denumerators), which places the values of $P'$ between zero and one.

We see at once that this is the special case of the last subsection, in which a single simple random variable is constrained to have a certain posterior expectation, while relative information is minimized. But to say this is to ignore the important difference between Field's approach and the present one. For he conceives of the input representable by commands of the form "multiply your prior probability for A by factor $e^{b(A)}/Z$!" While I would object that this conception leaves no room for the justification question of how well, or badly, an agent acted in his response to experience (assuming that he obeys the command), I will briefly explore this alternative representation.

To begin let us note that a pure Jeffrey shift is identical to a change in the *odds* for members of a given partition. That is, for a finite partition $X$ on which $P$ is positive, and which is a $P$-base for $P'$, the following are strictly equivalent:

(1) There are positive numbers $q(A)$ summing to 1 such that $P'(A) = q(A)P(A)$,
(2) There are positive numbers $k(AB)$ such that $P'(A)/P'(B) = k(AB)P(A)/P(B)$, for all $A$ and $B$ in $X$.

For we can define $k(AB)$ as $q(A)/q(B)$; or conversely, we note that the $q(A)$ must sum to 1 and must be defined so that

$$q(A)P(A) = k(AB)\,q(B)P(B)[P(A)/P(B)]$$

hence

$$q(A) = k(AB)q(B)P(A).$$

So $1 = \Sigma\{k(AB)q(B)P(A)\text{: } A \text{ in } X\}$ which means that

$$q(B) = 1/\Sigma\{k(AB)P(A)\text{: } A \text{ in } X\}.$$

Thus there is no doubt that each pure Jeffrey shift can be interpreted equivalently as obeying a command of form "Redistribute your odds on partition $X$ by multiplying them by these given coefficients!" But (2) is in turn equivalent to

(3) There are real numbers $b(A)$ summing to zero such that $P'(A) = P(A)e^{b(A)}/\Sigma\{P(B)e^{b(B)}: B \text{ in } X\}$.

In one direction, this equivalence is obvious, since $k(AB)$ can be defined as from the $q(A)$. Conversely, given the numbers $k(AB)$ we can find the numbers $b(A)$ as follows. They must satisfy $\log(k(AB)) = \log(e^{b(A)}/e^{b(B)})$ which equals $b(A) - b(B)$. So, summing over $B$ in $X$, we get

$$\Sigma\{\log k(AB): A \text{ in } X\} = -nb(B)$$

where $n$ is the cardinality of $X$ and I used the fact that the $b(A)$ must sum to zero. These reflections immediately yield a definition of $b(B)$ in terms of the coefficients $k(AB)$, and a short calculation shows that so defined, (3) does follow from (2) as well.

The conclusion is that an alternative form of epistemic input can be represented as a command to change the odds in a predetermined manner. Such commands, if obeyed, give rise to pure Jeffrey shifts; these can in turn be combined with prime shifts (in response to other sorts of epistemic input) so as to yield any Jeffrey shift.

**6. Epistemic States: A Second Theory.** Not everyone is inclined to accept the simple theory of the preceding three sections. Doubts are of two sorts. First, it is argued that simple conditionalization does occur and is in fact a form of acceptance, of the formation of full belief. But acceptance of a proposition $E$ as evidence is itself a fallible process, and later experience must therefore be able to lead us to revise the status of $E$. Neither simple nor Jeffrey conditionalization allows for this: if $P(E)$ is one, no Jeffrey shift will lead to a lower probability for it.

The second doubt concerns the process of revision on the basis of evidence which is compatible with all that has been accepted so far. Traditionally, a proposition may be newly accepted as true in accordance with some inductive procedure. For example, a hypothesis may be accepted because it is, of a contemplated range of hypotheses, the best explanation of the accepted evidence; or rejected because it fares so badly as an explanation in comparison with the others. Orthodox statistical testing is generally described similarly; inference to the best explanation is a rule often cited and discussed in general philosophy of science. Such procedures conflict, at least at first sight, with conditionalization, for something appears to be accepted which goes beyond the evidence.

In response to these doubts, epistemic states and inputs may be conceived more widely. I shall describe, first intuitively and then precisely, such a more general conception. To begin with the state: the person has a body of propositions, his *(rational) corpus K*; these are the ones he fully believes, accepts, is willing to use as evidence, what have you. Secondly, he has some function, I shall call it his *recipe,* which determines what epistemic judgments he is committed to, on the basis of what his corpus is: these are his *accepted epistemic judgments*. The *admissible valuations* are exactly like those of the first theory, namely probability functions on the underlying field *F* of factual propositions. His accepted epistemic judgments determine a subclass of these, the *class of representing probabilities;* namely, just those valuations which satisfy all the accepted judgments.

Both the corpus and the recipe can change with time. In the case of the recipe, there may be a number of further variables on which it depends, such as the range of contemplated hypotheses, or a reliability rating on the propositions in the corpus (reflecting how well supported their membership in the corpus is). Thus we must admit as epistemic input at least two new sorts of commands: to change the corpus in some fashion, or to change (some variable determining) the recipe.

Levi (1974) and Harper (1977) both have discussions of belief change which suggest here a dynamics for change in the corpus. (I emphasize that I am only adapting their remarks and not attempting to describe their theories.) An input takes the form "Accept proposition *E* into your corpus!" If *E* is consistent with the prior corpus, it is simply added to form the posterior corpus. If not, then the corpus must be revised in a way that satisfies the constraint (*E* belongs to the posterior corpus) and is optimal in other respects (the change is in some sense minimal, among those changes that satisfy the constraint).

An interesting question here is whether we should have feedback: if we subsequently accept hypothesis *H* in accordance with some inductive practice, such as inference to the best explanation, does that mean merely that we become committed to the epistemic judgment expressing full belief in *H,* or also that *H* is added to the corpus? If the latter is the case, simple conditionalization must, apparently, be violated by any such inductive practice.

The preceding two sections concern the *epistemic commitment,* that is, the function which governs (at least, as far as the person's present intention goes) changes in the state in response to the deliverances of experience. It does not seem to me that the sorts of input discussed previously (Section 4) become irrelevant: they yield constraints on the posterior class of representing probabilities, and so indirectly, on the remainder of the posterior state. Conversely, the new inputs impose constraints indirectly on the posterior representing probabilities. It is this relation that allows, presumably, some transfer from the earlier discussion to this wider context. (The first theory is a special case of the second theory, if we identify for example the corpus with the set of propositions that receive value one, and so forth. Below I shall suggest that *in addition* the first sort of theory may be regarded from the second point of view, as a description of the "surface phenomena.")

It may be clear that I have tried to make room in this scheme also for the very disparate approaches of Giere (1975), Harman (1965), Harper (1977), Salmon (1977), and Kyburg (1974). Since the scheme is still relatively simple it would surely prove Procrustean if we tried actually to fit their theories into its mold. One major difference among these writers, not touched on here, concerns what sorts of propositions go into the field of "factual" propositions, which furnishes the members of the corpus. Will these include statements of objective chance, or only of statistical distribution and relative frequency? The latter alternative is espoused in different ways by Kyburg and Salmon; the former, also in very different ways by Giere and Levi. For the latter, Kyburg has explicitly given what I call here a recipe (his epistemic judgments all being of the form "The probability that $A$ is no less than $r$, no greater than $s$") which depends on no hidden variables. For the case of objective chance, David Lewis (1978) and I (1980) have given recipes that generalize the idea that my personal conditional probability for $A$, given that the chance of $A$ equals $r$, equals $r$. (I have learned from Jeffrey that this principle was originally stated by Hacking; in (1980) I referred to it as the basic premise of Miller's paradox.)

**7. Representation of These States.** There is a field $F$ of "factual" propositions. A *corpus* is a set of such propositions. Criteria of rationality may delimit among these the rational corpora through such conditions as joint satisfiability or closure under entailment. Epistemic judgments take the four forms described in the typology of Section 2. Admissible epistemic valuations are probability functions defined on field $F$.

A *recipe* is a function that maps each corpus into a set of epistemic judgments, the *accepted epistemic judgments*. Again criteria of rationality may be applied; for example, if $A$ is in the corpus, the judgment expressing full belief in $A$ should be among the accepted judgments, and if the corpus is a satisfiable set of factual propositions, then the accepted judgments must be jointly satisfiable by some admissible epistemic valuation.

Suppose the corpus is $K$, the recipe $C$, and hence the set of accepted judgments, $C(K)$. Then the set of admissible valuations (i.e. probability functions on $F$) is the *representor class* or *class of representing probabilities,* and I shall denote it $C(K)^*$. The quadruple consisting of $K$, $C$, $C(K)$, and $C(K)^*$ is the epistemic state.

Because the representation of epistemic states which I have just described bears an unmistakable resemblance to features of the epistemological theories of Kyburg and Levi, we must consider the question whether the representor class is convex. That depends on the chosen typology of epistemic judgments. In the case of the one I chose, that is indeed so; and it remains so for various extensions thereof.

For suppose that $P$ and $P'$ satisfy one of those epistemic judgments: $P(A) \leq P(B)$ and $P'(A) \leq P'(B)$ for example. In that case $eP(A) + (1 - e)P'(A) \leq eP(B) + (1 - e)P'(B)$, where $e$ lies strictly between zero and one. Similarly for the *greater than* relationship and for membership in an interval.

Indeed, if we extended epistemic judgments to personal conditional probabilities or personal odds, the representor class would still be convex. If $x/y \leq r$ and $x'/y' \leq r$ then $(ex + (1 - e)x')/(ey + (1 - e)y')$ is also less than or equal to $r$. It is only if we start adding judgments of the form "The probability, to me, that $A$, equals either $r$ or $s$" or more generally "The probability, to me, that $A$, lies in Borel class $E$" that this convexity disappears.

This will explain why convex classes of probability functions have appeared at this point in epistemological theory. But I still think it is too much an accident of typology to be a principle.

As to input, constraints on the posterior probability of the sorts discussed in Section 4 appear here as constraints on the class of representing probabilities of the posterior state. Thus, the (self-)command "Change the probability for $A$ to $q$!" is satisfied if $P(A) = q$ for all $P$ in the class of all probability functions that satisfy the accepted epistemic judgment. So the command is to change the epistemic state in such a way that this happens. However, in this theory that can be done only by changing the rational corpus and/or recipe. Such changes can be made in a way as to satisfy the new constraint, and we may be able to evaluate the rationality of the way actually adopted. But clearly we should admit also new sorts of input, which relate directly to the corpus and to the parameters on which the recipe depends "admit $E$ into your corpus!" and "admit

$H$ into your range of contemplated hypotheses!'' are two such new sorts of input.

The new input has the indirect effect of constraining the class of representing probabilities as well; but only relative to the epistemic commitment which governs the reaction to that input. This effect is an *induced constraint* on the representing probabilities; for example, if $E$ is consistent with the corpus and the epistemic commitment holds the recipe constant when the input is the command ''admit $E$ into your corpus!'' then the satisfaction of that command will presumably constrain the posterior representing probabilities to assign 1 to $E$. It seems to me that Levi's critique of Jeffrey (Levi 1967) would suggest here that this is the only way in which the probability judgments concerning $E$ can rationally change. But the only reason for that which I can see is that otherwise, no criteria of rationality can be applied to the change; and *that* is not cogent: we can evaluate the manner in which the constraint is actually satisfied by comparing the ways in which it could be satisfied.

Since all constraints therefore produce induced restraints on the class of representing probabilities, we can think of probability kinematics here as the theory of ''surface phenomena,'' that is, of the changes in accepted epistemic judgments, equivalent in this theory (due to the typology of judgments adopted) to changes in the class of representing probability functions. I shall now turn to a simple instance of this kind of theory to provide a concrete case-study.

**8. The Statistical Syllogism and Jeffrey Conditionalization.** Let us suppose that factual propositions are generated by simple statements of two sorts: one attributes membership in certain classes to individuals, and the other states what proportion of members of one class belong to another class. The recipe has as paradigm the statistical syllogism, of which the following is a typical example:

(1) 90% of Swedes are Protestant;
(2) Petersen is a Swede;
(3) I have no information relevant to the question whether Petersen is a Protestant beyond 1 and 2 above.

Therefore: (4) the probability (to me) that Petersen is Protestant equals 0.9.

Reichenbach, Salmon, and Kyburg are especially associated with theories that will come to mind here. In order to emphasize the difference between this point of view and the subjective Bayesian one, it may be more apt to replace ''to me'' in (4) with ''relative to my total body of information.'' It may also be convenient to take a cue from Kyburg's terminology and rephrase (3) as "Petersen is (relative to my body of information) a random Swede with respect to the class of Protestants."

The great difficulty will of course be to generalize the recipe leading from corpus to epistemic judgments beyond this paradigm case. I shall take up two possibilities with respect to a somewhat more complex example, which was suggested by Levi's recent critique of Kyburg. Suppose that rational corpus $K$ contains exactly three propositions (except perhaps for ones that follow from these three):

*A*. Petersen is a Swede, and a resident of Malmö.
*B*. 90% of Swedes are Protestants.
(*C-or-D*). Either 85% or 95% of Swedes who are residents of Malmö are Protestant.

Whatever recipe is adopted, it should yield probability one judgments for each of these three propositions. I shall abbreviate these judgments to: *prob* $(A) = 1$, *prob* $(B) = 1$, and *prob* (*C-or-D*) $= 1$, and say that they *hold* in the epistemic state which has corpus $K$ and such a recipe.

What now of the probability that Petersen is Protestant? Kyburg has advocated a point of view, according to which one can argue as follows. The proportion of Protestants among Swedish residents of Malmö is not less than 85%, no greater than 95%. That gives us simply less precise, and not conflicting, information concerning the question whether Petersen is Protestant, than $A$ and $B$ do together. It is best to use the most precise information we have which is not obviated by other information. Hence the correct recipe yields here the judgment that the probability that Petersen is Protestant is 0.90. Let us call this *Policy I.* (Please note that this is not meant as a description of Kyburg's much more sophisticated theory, but only takes up for our simple case the sort of suggestions he makes in a general way.)

A different line of thought is suggested by Salmon's notion of *homogeneous reference classes.*[5] Trying to follow that idea, I would argue here: because of (*C-or-D*), the class of Swedes is specifically believed not to be homogeneous with respect to Protestantism, that is, it has sub-classes which are explicitly believed to contain *given* different proportions of Protestants than the whole class does. But as basis for our epistemic judgments we should choose the widest reference classes not inhomogeneous relative to our information in

[5]This may not correspond to Salmon's use of the idea, especially as he concentrates on objective rather than epistemic homogeneity.

this sense. (Note the emphasis on "*given*": we know that the proportion of Protestants in the unit class of Petersen is either zero or one on merely general grounds. In a recent, unpublished manuscript Salmon deals with similar difficulties in a way I won't try to reproduce (Salmon 1977).) So in this case, and within the limits of our chosen typology of epistemic judgments, we are led to the following as basic accepted epistemic judgments to be yielded by the recipe: the probability that Petersen is a Protestant is no less than 0.85, no greater than 0.95. Let us call this *Policy II.*

The classes of representing probabilities are of course different in the two cases; let us call them *QI* and *QII respectively:*

(a) $P$ is in $QI$ iff $P(A) = P(B) = P(C\text{-}or\text{-}D) = 1$ and $P(E) = 0.9$
(b) $P$ is in $QII$ iff $P(A) = P(B) = P(C\text{-}or\text{-}D) = 1$ and $P(E)$ is in $[0.85, 0.95]$.

Where "$E$" abbreviates "Petersen is a Protestant."

We now consider two possible scenarios. In the first, corpus $K$ is augmented with proposition $C$, in the other with proposition $D$. We assume that under these circumstances, the epistemic commitment holds the recipe the same. Not only that, since $C$ and $D$ provide exact proportions of Protestants in Swedish Malmö, a subclass of the Swedes, proposition $B$ becomes irrelevant, and the recipe performs a paradigm instance of the statistical syllogism. The probability of Petersen being a Protestant changes in the one scenario to 0.85, in the other to 0.95. The posterior classes of representing probabilities are, independently of the policy adopted:

First scenario: $P$ is in $Q'$ iff $P(A) = P(B) = P(C) = 1$ and $P(E) = 0.85$
Second scenario: $P$ is in $Q''$ iff $P(A) = P(B) = P(D) = 1$ and $P(E) = 0.95$.

Note that these define the two posterior representor classes. Taking "$Q$" as a variable ranging over $QI$ and $QII$ only, the principle of simple conditionalization as applied to the representing probabilities would here have the form:

(CL) (a) If $P(C)$ is positive, and $P$ belongs to $Q$ then $P_C$ belongs to $Q'$
(b) If $P'$ belongs to $Q'$ then $P' = P_C$ for some $P$ in $Q$
(c) If $P(D)$ is positive, and $P$ belongs to $Q$ then $P_D$ belongs to $Q''$.
(d) If $P''$ belongs to $Q''$, then $P'' = P_D$ for some $P$ in $Q$.

If Policy II is followed, then (b) and (d) hold. For example, if $P'$ belongs to $Q'$, then it also belongs to $Q$, and $P'(C) = 1$ so $P' = P'_C$. But (a) and (c) do not hold, for there is no reason to think that if $P$ belongs to *QII* and assigns a positive value to $C$, that $P(E \mid C)$ will be 0.85. What it will be, depends on what $P(E\&C)$ and $P(C)$ are.

If Policy I is followed, none of the four clauses in (CL) holds. For (a) and (c) the reason is as above. For (b) and (d), imagine that they do hold, and that $P$ is in *QI*. Then we have:

$$P(C\text{-}or\text{-}D) = 1$$

$$P(E \mid C) = 0.85$$

$$P(E \mid D) = 0.95$$

$$\begin{aligned} P(E) &= 0.9 = P(E \mid C)P(C \mid C\text{-}or\text{-}D) + P(E \mid D)P(D \mid C\text{-}or\text{-}D) \\ &= P(E \mid C)P(C) + P(E \mid D)P(D) \\ &= 0.85\ P(C) + 0.95\ P(D) \\ &= 0.85\ P(C) + 0.95\ (1 - P(C)) \\ &= 0.95 - 0.1\ P(C) \end{aligned}$$

so that $P(C) = 1/2$ (where the calculations assume that $C$ and $D$ are disjoint). But this is in contradiction with the definition of *QI*, which makes it a class containing many functions $P$ that do not assign 1/2 to $C$. Some of these will violate change (b) and some clause (d).

I should like to point out that (CL) would fare better yet under Policy II if I had not chosen such a small typology of epistemic judgments. If the typology admitted judgments of the sort "The probability to me of $E$ is in set $R$," where $R$ can be *any* Borel set of real numbers within the unit interval, and not merely an interval, then the class *QII* would have been different: all its members would have assigned either 0.85 or 0.95 to $C$. If in addition, we added epistemic conditional judgments so as to express the independence of $E$ and $C$, and also of $E$ and $D$—which may seem reasonable here—then clauses (a) and (c) would also have held. Hence principle (CL), which adapts simple conditionalization could have held for the Salmon-like Policy II, if it were not for the restricted form of our typology of judgments.

Let us now go on to adapt similarly the rule of Jeffrey conditionalization to this case. Classes *QI, QII, Q, Q', Q''* are all as before.

(JL) There are Jeffrey shifts $f$ and $g$ such that

(i) if $P$ is in $Q$, and $P(C)$ is positive, then $fP$ is in $Q'$ and if $P(D)$ is positive, $gP$ is in $Q''$;
(ii) if $P'$ is in $Q'$ then $P' = fP$ for some $P$ in $Q$; and if $P''$ is in $Q''$ then $P'' = gP$ for some $P$ in $Q$.

That (JL) is correct in all respects is easy to establish. For Policy I, choose as base for $f$ the partition $\{C \cap E, C \cap \textit{not-}E, \textit{not-}C\}$. Define $f$ to be the shift:

$$fP(Y) = 0.85\,P(Y|C \cap E) + 0.15\,P(Y|C \cap \textit{not-}E) + 0P(Y|\textit{not-}C)$$

This is not a pure shift, but is a Jeffrey shift. We note that if $P$ is in $QI$ then $fP$ will assign *one* to $A$, $B$, and $C$, and that $fP(E) = 0.85$ as required, provided $P(C)$ was positive. Secondly, if $P'$ is in $Q'$, we have $P'(A) = P'(B) = P'(C) = 1$ and $P'(E) = 0.85$. We find $P$ as follows:

$$P(Y) = 0.9\,P'(Y|E \cap C) + 0.1\,P'(Y|\textit{not-}E \cap C)$$

The function $P$ so defined assigns 0.9 to $E$, and 1 to $A$, $B$, ($C$-*or*-$D$); hence belongs to $QI$. But also,

$$\begin{aligned} fP(Y) &= 0.85\,P(Y|C \cap E) + 0.15\,P(Y|C \cap \textit{not-}E) \\ &= 0.85\,P'(Y|C \cap E) + 0.15\,P'(Y|C \cap \textit{not-}E) \\ &= P'(Y) \end{aligned}$$

for since $P(C) = P'(C) = 1$, the conjunction with $C$ makes no difference; and $P$ was defined so that $P(-|E) = P'(-|E)$; and $P'(E) = 0.85$.

So (JL) is correct if Policy I was adopted; for of course $g$ can be defined similarly. In the case of Policy II exactly the same argument works. For in the argument for clause (i) the fact that $P(E) = 0.9$ if $P$ is in $QI$ is needed only to ensure that $P(E)$ is positive. For clause (ii), if we define $f$ exactly as above, we find a function in $QI$, which is a subclass of $QII$.

The conclusion is therefore that whether Policy I or II is chosen, and even while sticking to our smallish typology of epistemic judgments, Jeffrey conditionalization (as adapted here) correctly describes the change from prior to posterior representing probabilities. Although the example considered is still relatively simple, and the form of theory also, I consider this good evidence that Jeffrey conditionalization is compatible with a range of epistemological theories that extends well beyond the one for which it was originally designed.

REFERENCES

Domotor, Z., Zanotti, M. and Graves, H. (1979), "Probability Kinematics," (forthcoming) *Philosophy of Science 47.*

Field, H. (1978), "A Note on Jeffrey Conditionalization," *Philosophy of Science 45:* 361–367.

Giere, R. (1975), "The Epistemological Roots of Scientific Knowledge," in *Minnesota Studies in the Philosophy of Science* VI.

Hacking, I. (1967), "Slightly More Realistic Personal Probability," *Philosophy of Science 34*: 311–325.

Harman, G. (1965), "The Inference to the Best Explanation," *Philosophical Review 74*: 88–95.

Harper, W. (1977), "Rational Conceptual Change," in F. Suppe and P. Asquith (eds.) *PSA 1976, volume II.*

Hobson, A. (1971), *Concepts in Statistical Mechanics.* New York: Gordon and Breach.

Jamison, B. (1974), "A Martin Boundary Interpretation of the Maximum Entropy Argument" *Z. Wahrscheinlichkeitstheorie verw. Gebiete 30*: 265–272.

Jeffrey, R. (1965), *The Logic of Decision.* New York: McGraw Hill.

Kyburg, H. (1974), *The Logical Foundations of Statistical Inference.* Dordrecht: Reidel.

Kyburg, H. (1977), "Randomness and the Right Reference Class," *Journal of Philosophy 74*: 501–521.

Levi, I. (1967), "Probability Kinematics," *British Journal for Philosophy of Science 18*: 197–209.

Levi, I. (1974), "On Indeterminate Probabilities," *Journal of Philosophy 71*: 391–418.

Levi, I. (1977), "Direct Inference," *Journal of Philosophy 74*: 5–29.

Lewis, D. (1978), "A Subjectivist's Guide to Objective Chance," circulated ms.

May, S. and Harper W. (1976), "Toward an Optimization Procedure for applying Minimum Change Principles in Probability Kinematics," in W. L. Harper and C. A. Hooker (eds.) *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science, Volume I.* Dordrecht: Reidel.

Salmon, W. (1977), "Objectively Homogeneous Reference Classes" *Synthese 36*: 339–414. (A revision to be found in a forthcoming book.)

van Fraassen, B. (1979), "Foundations of Probability Theory: A Modal Frequency Interpretation," in G. Toraldo di Francia (ed.) *Problems in the Foundations of Physics.* Amsterdam: North-Holland.

van Fraassen, B. (1980), "A Temporal Framework for Conditionals and Chance," *Philosophical Review 89:* 91–108.

Williams, P. M. (1978), "Bayesian Conditionalization and the Principle of Minimum Information," presented *British Society for Philosophy of Science*; forthcoming.