Jeffrey Conditioning and External Bayesianity July 2009

Carl Wagner
Department of Mathematics
The University of Tennessee
Knoxville, TN 37996-1300, USA
wagner@math.utk.edu

Abstract. Suppose that several individuals who have separately assessed prior probability distributions over a set of possible states of the world wish to pool their individual distributions into a single group distribution, while taking into account jointly perceived new evidence. They have the option of (i) first updating their individual priors and then pooling the resulting posteriors or (ii) first pooling their priors and then updating the resulting group prior. If the pooling method that they employ is such that they arrive at the same final distribution in both cases, the method is said to be *externally Bayesian*, a property first studied by Madansky (1964). We show that a pooling method for discrete distributions is externally Bayesian if and only if it commutes with Jeffrey conditioning, parameterized in terms of certain ratios of new to old odds, as in Wagner (2002), rather than in terms of the posterior probabilities of members of the disjoint family of events on which such conditioning originates.

1. Combining Probability Distributions.

In what follows, Ω denotes a countable set of possible states of the world, assumed to be mutually exclusive and exhaustive. A function $p: \Omega \rightarrow [0,1]$ is a *probability mass function* (pmf) iff $\sum_{\omega \in \Omega} p(\omega) = 1$. The *support* of a pmf p is the set $\text{Supp}(p) := \{ \omega : p(\omega) > 0 \}$. Each pmf p gives rise to a *probability measure* (which, abusing notation, we also denote by p) defined for each set $E \subseteq \Omega$ by $p(E) := \sum_{\omega \in E} p(\omega)$.

Denote by Δ the set of all pmfs on Ω, and let n be a positive integer. Let Δ^n denote the n-fold Cartesian product of Δ, with

$$\Delta^{n+} := \{ (p_1,\ldots,p_n) \in \Delta^n : \cap_i \text{Supp}(p_i) \neq \varnothing \} .$$

A *pooling operator* is any function $T: \Delta^{n+} \rightarrow \Delta$.[1] Given $(p_1,\ldots,p_n) \in \Delta^{n+}$, the pmf $T(p_1,\ldots,p_n)$ may, depending on the context, represent

(i) a rough summary of the current pmfs $p_1,\ldots,p_n$ of n individuals;

(ii) a compromise adopted by these individuals in order to complete an exercise in group decision making;

(iii) a "rational" consensus to which all individuals have revised their initial pmfs $p_1,\ldots,p_n$ after extensive discussion;

(iv) the pmf of a decision maker external to a group of n experts (who may or may not have assessed his own prior over Ω before consulting the group) upon being apprised of the pmfs $p_1,\ldots,p_n$ of these experts;

(v) a revision of the pmf p_i of a particular individual i upon being apprised of the pmfs of individuals 1,…,i-1,i+1,…,n, each of whom he may or may not consider to be his "epistemic peer."

The appropriate restrictions to place on the pooling operator T will naturally depend on the interpretation of $T(p_1,\ldots,p_n)$. For example, it might seem reasonable for T to *preserve unanimity* $(T(p,\ldots,p) = p)$ [2] in cases (i),(ii), and (v) above, but perhaps not in cases (iii) and (iv). There is an extensive literature on this general subject (see, for example, the article of Genest and Zidek (1986) for a summary and appraisal of work done through the mid-1980s). There has also been a recent surge of interest in problems associated with interpretation (v) above, as part of a field of inquiry that has come to be termed the "epistemology of disagreement."

Our interest here is not in adjudicating which restrictions on pooling are appropriate in which situations, but rather in a formal analysis of one proposed group rationality condition, known as *external Bayesianity*, and its connections with Jeffrey conditioning.(We employ in what follows the language of interpretations (ii) and (iii) above, since it is in those cases where external Bayesianity seems most compelling. But our formal results apply to the other interpretations as well, for whatever interest that may have.) We shall see that a pooling operator for pmfs is externally Bayesian if and only if it commutes with Jeffrey conditioning, parameterized in terms of certain ratios of new to old odds, as in Wagner (2002), rather than in terms of the posterior probabilities of members of the disjoint family of events on which such conditioning originates.

2. Externally Bayesian Pooling Operators.

Consider the situation in which n individuals who have assessed pmfs $p_1,\ldots,p_n$ over Ω subsequently undergo identical new learning as a result of jointly perceived new evidence. Should they first update their individual priors based on the jointly perceived new information, and then pool the posteriors? Or should they pool their priors, and then update the result of pooling based on this information?

Take the simplest case, where each individual comes to learn that the true state of the world belongs to the subset E of Ω, but nothing that would change the odds between any states ω_1 and ω_2 in E. It would thus be appropriate for each individual i to revise his p_i to, let us call it q_i, by conditioning on E,[3] so that for each $\omega \in \Omega$,

(2.1) $$q_i(\omega) = p_i(\omega|E) := p_i(\omega)\,[\omega \in E] \,/\, p_i(E),$$

where $[\omega \in E]$ denotes the characteristic function[4] of the set E, evaluated at ω . These revised pmfs might then be combined by means of the pooling operator T. Alternatively, one might imagine first pooling the priors $p_1,\ldots,p_n$ and then conditioning the result on E. If either of these procedures results in the same final distribution, we say that T *commutes with conditioning* (CC). This property may be expressed formally as follows:

CC : For all subsets E of Ω and all $(p_1,\ldots,p_n) \in \Delta^{n+}$ such that $p_i(E) > 0$, $i= 1,\ldots,n$, and $(p_1(.|E),\ldots,p_n(.|E)) \in \Delta^{n+}$, it is the case that

(2.2) $$T(p_1,\ldots,p_n)(E) > 0 ,$$

and

(2.3) $$T(p_1(.|E),\ldots,p_n(.|E)) = T(p_1,\ldots,p_n)(.|E).$$

Condition CC is a special case of a more general property of pooling operators called *external Bayesianity*, first introduced by Madansky (1964,1978), and elaborated upon by Genest, McConway, and

Schervish (1986). Given $(p_1,\dots,p_n) \in \Delta^{n+}$, we say that a function $\lambda : \Omega \to [0, \infty)$ is a *likelihood* for $(p_1,\dots,p_n)$ if

(2.4) $$0 < \sum_{\omega\in\Omega} \lambda(\omega)\, p_i(\omega) < \infty \quad, i= 1,\dots,n,$$

and $(q_1,\dots,q_n) \in \Delta^{n+}$, where

(2.5) $$q_i(\omega) := \lambda(\omega)p_i(\omega) / \sum_{\omega\in\Omega} \lambda(\omega)\, p_i(\omega)\ .$$

For concreteness, it may be useful to think of the special case of the above in which the domain of the pmfs p_i is extended to include the event *obs*, with $q_i(\omega) = p_i(\omega|obs)$, and $\lambda(\omega)$ denoting the common value of the likelihoods $p_i(obs|\omega)$.[5] Then (2.5) is simply Bayes' formula, and models the revision of the priors $p_1,\dots,p_n$ in a situation where individuals disagree about the probabilities of various possible states of the world, while agreeing about the probability of experiencing *obs* in each of those states.[6]

A pooling operator $T : \Delta^{n+} \to \Delta$ is *externally Bayesian* (EB) iff the following condition is satisfied:

EB : If $(p_1,\dots,p_n) \in \Delta^{n+}$ and λ is a likelihood for $(p_1,\dots,p_n)$, then

(2.6) $$0 < \sum_{\omega\in\Omega} \lambda(\omega)\, T(p_1,\dots,p_n)(\omega) < \infty,$$

and the following commutativity property holds:

(2.7) $$T(\lambda p_1 / \sum_{\omega\in\Omega} \lambda(\omega)\, p_1(\omega), \dots , \lambda p_n / \sum_{\omega\in\Omega} \lambda(\omega)\, p_n(\omega))$$

$$= \lambda\, T(p_1,\dots,p_n) / \sum_{\omega\in\Omega} \lambda(\omega)\, T(p_1,\dots,p_n)(\omega).$$ [7]

Implicit in the definition of external Bayesianity is the following view of the proper way to represent identical new learning: Recall that if q is a revision of the probability measure p and A and B are events, the *Bayes factor* $\beta(q,p ; A : B)$ is the ratio

(2.8) $$\beta(q,p ; A:B) := (q(A)/q(B)) / (p(A)/p(B))$$

of new to old odds, and the *relevance quotient* $\rho(q,p ; A)$ is the ratio

(2.9) $$\rho(q,p ; A) := q(A)/p(A)$$

of new to old probabilities. When $q = p(.|E)$, then (2.8) is simply the *likelihood ratio* $p(E|A)/p(E|B)$. More generally,

(2.10) $$\beta(q,p\ ;\ A : B) = \rho(q,p\ ;\ A) / \rho(q,p\ ;\ B)\ ,$$

a simple, but useful, identity. Suppose that the probability distributions $(q_1,\ldots,q_n)$ are related by formula (2.5) to the distributions $(p_1,\ldots,p_n)$. Then, for all i, and for all ω and ω* in Ω, it is easy to verify that

(2.11) $$\beta(q_i,p_i\ ;\ \omega : \omega^*) = \lambda(\omega)/\ \lambda(\omega^*),$$ [8]

i.e., for each fixed pair of states in Ω, the ratios of new to old odds are *uniform* over i = 1,…,n. Similarly, denoting $T(p_1,\ldots,p_n)$ by p and $\lambda\, T(p_1,\ldots,p_n) / \sum_{\omega\in\Omega} \lambda(\omega)\, T(p_1,\ldots,p_n)(\omega)$ by r, we have

(2.12) $$\beta(r,p\ ;\ \omega : \omega^*) = \lambda(\omega)/\ \lambda(\omega^*).$$

In other words, implicit in the definition of external Bayesianity is the view that, *identical new learning should be reflected in identical Bayes factors at the level of atomic events.* This is an important point, to which we shall have occasion to return in Section 3.

Remark 2.1. Setting $\lambda(\omega) = [\omega \in E]$ shows, as suggested above, that EB implies CC.

Remark 2.2. The term *externally Bayesian* derives from the fact that a group of decision makers having a common utility function but different priors over the relevant states of nature and employing an EB pooling operator will make decisions that appear to an outsider like the decisions of a single Bayesian decision maker (see Madansky 1964). An example due to Raiffa (1968, pp. 221-6) shows that the use of pooling operators that fail to satisfy EB may lead members of the group to act in strange ways.

The set of externally Bayesian pooling operators is clearly nonempty since “dictatorial” pooling (for fixed d, $T(p_1,\ldots,p_n) = p_d$ for all $(p_1,\ldots,p_n) \in \Delta^{n+}$) satisfies EB. Note, however, that weighted arithmetic means, i.e., pooling operators T defined by

(2.13) $$T(p_1,\ldots,p_n) := \sum_{1\le i \le n} w(i)p_i(\omega),$$

where w(1), … ,w(n) is a sequence of nonnegative real numbers summing to one, fail in general to satisfy EB. On the other hand, suppose that we define a pooling operator T by

(2.14) $T(p_1,\ldots,p_n)(\omega) := \Pi_{1\leq i \leq n}\, p_i(\omega)^{w(i)} / \sum_{\omega\in \Omega} \Pi_{1\leq i \leq n}\, p_i(\omega)^{w(i)}$, [9]

where $0^0 := 1$. It is easy to verify that such normalized weighted *geometric* means, which have come to be termed *logarithmic pooling operators*, are externally Bayesian, a fact first noted (according to Bacharach (1972)) by Peter Hammond. A complete characterization of externally Bayesian pooling operators appears in Genest, McConway, and Schervish (1986).

3. Jeffrey Conditioning and External Bayesianity.

Let Ω be a set of possible states of the world (with no cardinality restrictions for the moment) and let **A** be a σ-algebra of subsets ("events") of Ω. Let p and q be probability measures on **A**, and let **E** = $\{E_k\}$ be a countable set of nonempty, pairwise disjoint events such that $p(E_k) > 0$ for all k. We say that q comes from p by *Jeffrey conditioning* (JC) on **E** if there exists a sequence (e_k) of positive real numbers summing to one such that, for every A in **A**,

(3.1) $q(A) = \sum_k e_k\, p(A|E_k)$. [10]

Formula (3.1) is the appropriate way to update your prior p in light of new evidence if and only if (1) based on the total evidence, old as well as new, you judge that for each k, the posterior probability $q(E_k)$ should take the value e_k ; and (2) for each E_k, and for all A in **A**, you judge that $q(A|E_k) = p(A|E_k)$.[11] Conditions (1) and (2) amount to the assertion that you have learned something new about the events E_k, but nothing new about their relevance to any other events. Of course, when **E** = {E}, JC reduces to ordinary conditioning.

In the special case where Ω is countable and **A** = 2^{Ω}, formula (3.1) is equivalent to

(3.2) $q(\omega) = \sum_k e_k\, p(\omega|E_k) = \sum_k e_k\, p(\omega)\, [\omega \in E_k] / p(E_k)$, for all $\omega\in \Omega$, and condition (2) is equivalent to the judgment that, for all k, nothing in the new evidence should disturb the odds between any two states of the world in E_k.[12] Note also that if p and q are *any* pmfs on the countable set Ω and $\mathrm{Supp}(q) \subseteq \mathrm{Supp}(p)$, then q comes from p by JC on the family $\mathbf{E} = \{ \{\omega\} : q(\omega) > 0 \}$.

Does every externally Bayesian pooling operator commute with Jeffrey, as well as ordinary, conditioning? Consider the following simple example, where $\Omega = \{ \omega_1, \omega_2, \omega_3, \omega_4 \}$, $E_1 = \{ \omega_1, \omega_2 \}$, $E_2 = \{ \omega_3, \omega_4 \}$, $\mathbf{E} = \{ E_1, E_2 \}$, and $e_1 = e_2 = ½$:

TABLE 3.1

	ω_1	ω_2	ω_3	ω_4			ω_1	ω_2	ω_3	ω_4
p_1 :	1/7	4/7	1/7	1/7	POOL →	p :	1/4	1/4	1/4	1/4
p_2 :	4/13	1/13	4/13	1/13				↓ JC		
		↓ JC				r :	1/4	1/4	1/4	1/4
q_1 :	1/10	2/5	1/4	1/4	POOL →	q :	2/9	2/9	5/18	5/18
q_2 :	2/5	1/10	1/4	1/4						

In the above revision schema, q_i comes from p_i by JC on the partition **E**, with $q_i(E_1) = q_i(E_2) = ½$, i = 1, 2, and r comes from p by JC on **E**, with $r(E_1) = r(E_2) = ½$. The distribution p results from the externally Bayesian pooling of p_1 and p_2 by normalized geometric averaging (i.e., by formula (2.14) above, with n =2 and w(1) = w(2) = ½), and the distribution q results from the same pooling formula, applied to q_1 and q_2. Since $q \neq r$, we apparently have here a case where updating by JC does not commute with externally Bayesian pooling. But this is neither surprising nor disturbing. For the Jeffrey revisions of p_1 to q_1, of p_2 to q_2, and of p to r, parameterized as in formula (3.2), are not instances of updating by a single likelihood λ. Contrary to (2.11) and (2.12), we have, for example, $\beta(q_1,p_1 ; \omega_1 : \omega_3) = 2/5$, $\beta(q_2,p_2 ; \omega_1 : \omega_3) = 8/5$, and $\beta(r,p ; \omega_1 : \omega_3) = 1$. On the conception of identical new learning implicit in the definition of external Bayesianity, agents 1 and 2 are *not* revising their priors p_1 and p_2 based on identical new learning. And their fictional group surrogate is revising the result, r, of pooling p_1 and p_2 in response to still different new learning. In short, in the context of updating by Jeffrey conditioning, identical new learning is *not* expressed by identical posterior probabilities of the events E_k.

What is the proper representation of identical new learning in this context? We have already observed that it is implicit in the definition of external Bayesianity that identical new learning should be reflected in identical Bayes factors at the level of atomic events. When updating by JC, however, there is an equivalent, coarser-grained version of this identity, based on the following special case of a theorem in Wagner (2003, Theorem 3.1):

Theorem 3.1. Given a countable set Ω, let p, q, P, and Q be probability measures on $\mathbf{A} = 2^{\Omega}$, with q regarded as a revision of p, and Q a revision of P. If q comes from p by Jeffrey conditioning on the family **E**, then

(3.3) $\beta(Q,P\ ;\ \omega : \omega^*) = \beta(q,p\ ;\ \omega : \omega^*)$ for all $\omega, \omega^* \in \Omega$

if and only if Q comes from P by Jeffrey conditioning on **E** and

(3.4) $\beta(Q,P\ ;\ E_i : E_j) = \beta(q,p\ ;\ E_i : E_j)$ for all $E_i, E_j \in \mathbf{E}$.

This suggests that, for our present purposes, the proper parameterization of Jeffrey conditioning should involve Bayes factors associated with pairs of events in **E**. Such a parameterization, which first appears in Wagner (2002)[13], is easily derived from formula (3.2).

Theorem 3.2. Suppose that q comes from p by Jeffrey conditioning on the family $\mathbf{E} = \{E_k\}$ and let $b_k := \beta(q,p\ ;\ E_k : E_1)$, k = 1,2,… . Then, for all ω in Ω,

(3.5) $q(\omega) = \sum_k b_k\, p(\omega)\, [\omega \in E_k] \ /\ \sum_k b_k\, p(E_k)$.[14]

Proof. Dividing (3.2) by $1 = q(\Omega) = \sum_k p(E_k)\, q(E_k)/p(E_k)$, replacing e_k by $q(E_k)$, dividing the numerator and denominator of the resulting fraction by $q(E_1)/ p(E_1)$, and invoking (2.10) yields (3.5). □

The above parameterization of JC allows us to formulate a commutativity property with the proper relation to external Bayesianity. We say that a pooling operator T commutes with JC, so parameterized, (CJC) if and only if the following condition holds:

CJC : For all families **E** = $\{E_k\}$ of nonempty, pairwise disjoint subsets of Ω, all $(p_1,\ldots,p_n) \in \Delta^{n+}$ such that $p_i(E_k) > 0$ for all i and all k, and all sequences (b_k) of positive real numbers such that $b_1 = 1$ and

(3.6) $$\sum_k b_k\, p_i(E_k) < \infty \quad , i= 1,\ldots,n,$$

and such that $(q_1,\ldots,q_n) \in \Delta^{n+}$, where

(3.7) $$q_i(\omega) := \sum_k b_k\, p_i(\omega)[\omega \in E_k] \;/\; \sum_k b_k\, p_i(E_k) \,,$$

it is the case that

(3.8) $$0 < \sum_k b_k\, T(p_1,\ldots, p_n)(E_k) < \infty \,,$$

and

(3.9) $$T(\sum_k b_k\, p_1[\,.\!\in E_k] / \sum_k b_k\, p_1(E_k), \ldots , \sum_k b_k\, p_n[\,.\!\in E_k] / \sum_k b_k\, p_n(E_k))$$

$$= \sum_k b_k\, T(p_1,\ldots, p_n)[\,.\!\in E_k] \;/\; \sum_k b_k\, T(p_1,\ldots, p_n)(E_k) \,.$$

Theorem 3.3. A pooling operator T: $\Delta^{n+} \rightarrow \Delta$ is externally Bayesian if and only if it commutes with Jeffrey conditioning in the sense of CJC .

Proof. *Necessity*. Let T be externally Bayesian. Suppose that (3.6) holds and that $(q_1,\ldots,q_n) \in \Delta^{n+}$, where q_i is given by (3.7).
Let $\lambda(\omega) = \sum_k b_k [\omega \in E_k]$. Then, for i = 1,…,n ,

(3.10) $$\sum_{\omega\in \Omega} \lambda(\omega)\, p_i(\omega) = \sum_{\omega\in \Omega} p_i(\omega) \sum_k b_k [\omega \in E_k]$$

$$= \sum_k b_k \sum_{\omega\in \Omega} p_i(\omega)[\omega \in E_k] = \sum_k b_k\, p_i(E_k) \,,$$

and so by (3.6) and the fact that each of the terms $b_k\, p_i(E_k)$ is positive, λ is a likelihood for $(p_1,\ldots,p_n)$. By (2.6) and (2.7) it then follows that (3.8) and (3.9) hold.

Sufficiency. Suppose that T satisfies CJC and that λ is a likelihood for $(p_1,\ldots,p_n)$. Let $(\omega_1,\omega_2,\ldots)$ be a list of all those $\omega \in \Omega$ for which $\lambda(\omega) > 0$, and let **E** $= \{ E_1, E_2 ,\ldots\}$, where $E_i := \{ \omega_i \}$. Setting $b_k := \lambda(\omega_k)/ \lambda(\omega_1)$ for k = 1,2,…, we have each $b_k > 0$ and $b_1 = 1$. From the fact that λ is a likelihood for $(p_1,\ldots,p_n)$ it follows that (3.6) holds and that $(q_1,\ldots,q_n) \in \Delta^{n+}$, where $q_i(\omega)$ is given by (3.7). Hence, by CJC, (3.8) and (3.9) hold. But (3.8) implies (2.6), and (3.9) implies (2.7). □

4. Discussion

On a minimalist interpretation of the foregoing, we have simply established the mathematical fact that if we want Jeffrey conditioning to commute with externally Bayesian pooling operators, then the right parameterization of such conditioning is given by (3.5), rather than (3.2). And the heuristic employed in identifying (3.5) can be seen as a purely formal inference from the fact that the very scenario in which external Bayesianity is formulated dictates that individuals update their priors using *identical likelihoods*, and that the result of pooling those priors is updated using those same likelihoods. Since identical likelihoods entail identical Bayes factors at the level of atomic events, we need only observe (Theorem 3.1) that for Jeffrey updating, the identity of atomic level Bayes factors is equivalent to the identity of Bayes factors at the level of events in the family **E**, and we arrive at the doorstep of the parameterization (3.5), and hence of the formulation of CJC that ensures that EB implies CJC. As a bonus, we get the converse implication, and thus a particularly salient characterization of EB in terms of Jeffrey conditioning, a probability revision method with which formal epistemologists are well acquainted.

As we suggested in section 2, however, the parameterization of Jeffrey conditioning in terms of Bayes factors is more than a mere mathematical expedient. After all, the definition of external Bayesianity dictates that individuals update their priors using identical likelihoods (hence, identical Bayes factors) because EB is an attempt to model a feature of probability pooling in a context where individuals undergo *identical new learning.* And the view that what is learned from new evidence alone should be represented by ratios of new to old odds (or their logarithms) has a distinguished pedigree, going back to I.J. Good (1950, 1983) and promoted in recent years in the work of Richard

Jeffrey (1992a, 2004), among others. In addition to being supported by the results of the present paper, this view is buttressed by several other theorems. In Wagner (2002) it is shown that Jeffrey updating on a family **E**, followed by Jeffrey updating on a family **F**, produces the same final result as first updating on **F**, and then on **E**, as long as identical new learning is represented by identical Bayes factors. Indeed, under mild regularity conditions, such a representation is necessary to ensure the commutativity of Jeffrey updates. See also Lange (2000) and Osherson (2002). In Wagner (1997,1999, 2001) it is shown that two ways of generalizing Jeffrey's (1992b, 1995) solution to the old evidence problem to the case of uncertain old evidence and statistical new explanation coincide when identical new learning is represented by identical Bayes factors. In Wagner (2003) the aforementioned results are given a unified treatment, and detailed arguments are presented in support of the general principle ("the uniformity rule") that identical new learning should be represented by identical Bayes factors at the level of atomic events.[15,16]

Notes

1. Many treatments of the problem of combining probability distributions deal with pooling operators defined, slightly more generally, on Δ^n. For reasons that will soon be obvious, we are excluding consideration of cases where there is no state of the world ω to which all individuals assign positive probability. This is, however, considerably less stringent than the usual restriction posited in treatments of external Bayesianity, namely, that for all pmfs p_i under consideration, $Supp(p_i) = \Omega$.

2. This is a rather weak unanimity condition. One often sees in the literature a much stronger condition of this type postulated, namely, that for each $c \in [0,1]$, if $p_1(\omega) = \ldots = p_n(\omega) = c$, then $T(p_1,\ldots,p_n)(\omega) = c$. It isn't clear that pooling should be restricted in this way in any of the scenarios (i)-(v).

3. See Jeffrey [1992a, pp.117-119].

4. Specifically, $[\omega \in E] = 1$ if $\omega \in E$, and 0 if $\omega \in E^c$. This notation is a special case of the wonderful notational device introduced by Kenneth Iverson [1962] in his programming language APL, in which propositional functions are denoted simply by enclosing the

propositions in question in square brackets. See Knuth [1992] for an elaboration of how this notation simplifies the representation of various restricted sums.

5. As statisticians use the term, any constant multiple of the elementary likelihoods p_i(*obs*|ω) is also called a likelihood function. It is for this reason (and also the fact that in the continuous case, conditional density functions may take values outside [0,1]) that λ is not restricted to taking values in the unit interval. See Evans and Rosenthal [2004, pp. 281-283] for a particularly lucid discussion of the likelihood function.

6. It should be emphasized, however, that if Ω is infinite, there are cases of probability revision that cannot be conceptualized as the result of conditioning on some event *obs* in an extension of Ω. This observation is the essence of the famous *Superconditioning Criterion* of Diaconis and Zabell [1982, Theorem 2.1].

7. The property of external Bayesianity actually applies to pooling operators for any sort of density function. We have described here what this property amounts to in the discrete case. For the full formulation, including a characterization of externally Bayesian pooling operators in the most general setting, see Genest, McConway, and Schervish (1986).

8. When formula (2.11) involves division by zero, it is to be understood as asserting that $\lambda(\omega^*)p_i(\omega^*)q_i(\omega) = \lambda(\omega)p_i(\omega)q_i(\omega^*)$. A similar remark applies to formula (2.12).

9. Note that $\sum_{\omega\in\Omega} \prod_{1\leq i\leq n} p_i(\omega)^{w(i)}$ is strictly positive since $(p_1,\ldots,p_n) \in \Delta^{n+}$, and finite since $\prod_{1\leq i\leq n} p_i(\omega)^{w(i)} \leq \sum_{1\leq i\leq n} w(i)p_i(\omega)$ by the generalized arithmetic-geometric mean inequality. See Royden (1963, p. 94).

10. In the standard exposition of Jeffrey conditioning (also called *probability kinematics*) the family **E** = $\{E_k\}$ is taken to be a *partition* of Ω, so that, in addition to pairwise disjointness of the events E_k, one has $\bigcup_k E_k = \Omega$. Standardly, (e_k) is a sequence of *nonnegative* real numbers summing to one, and it is assumed that zeros are not raised, i.e., that $p(E_k) = 0$ implies that e_k (= $q(E_k)$) = 0. Finally, it is stipulated that $0 \times p(A|E_k) = 0$ if $p(E_k) = 0$, so that $e_k\, p(A|E_k)$ is well-defined even if

$p(A|E_k)$ isn't. Given the standard formulation, our family **E** simply comprises those E_k in the partition for which $e_k > 0$. Conversely, our format yields the standard one by associating to our family **E** = $\{E_1, E_2,\ldots\}$, if it fails to be a partition, the partition **E** = $\{E_0, E_1, E_2,\ldots\}$, where $E_0 := \Omega \setminus \cup_k E_k$, and setting $e_0 = 0$.

11. Condition (2) is known variously as the *rigidity condition* or the *sufficiency condition*.

12. See Jeffrey [1992, pp. 124-125].

13. This parameterization of JC is in the spirit of an earlier parameterization, due to Hartry Field (1978). Field's parameterization only applies, however, to JC originating in a finite family of events.

14. When p and q are probability measures on a sigma algebra **A** of subsets of a set Ω of arbitrary cardinality, this parameterization takes the form $q(A) = \Sigma_k b_k p(A \cap E_k) / \Sigma_k b_k p(E_k)$ for all $A \in$ **A**.

15. In particular, it is shown in section 5 of that paper how other measures of probability change, such as ratios or differences between posterior and prior probabilities, fail to meet three criteria that should clearly be satisfied by any representation of new learning.

16. Note that the problem of how best to represent new learning is distinct from the problem of how to measure the extent to which evidence E confirms hypothesis H. In particular, adopting atomic level Bayes factors as the solution to the former problem does not commit one to adopting the likelihood ratio $p(E|H)/p(E|H^c)$, or some monotone transformation thereof, as the solution to the latter. See Joyce (2004) for a spirited defense of the view that the multiplicity of measures of confirmation is something to be celebrated, each capturing a different and useful sense of evidential support.

References

Bacharach, Michael (1972), "Scientific Disagreement," Unpublished manuscript.

Diaconis, Persi, and Sandy Zabell, "Updating Subjective Probability," *Journal of the American Statistical Association* 77: 822-830.

Evans, Michael J., and Jeffrey S. Rosenthal (2004), *Probability and Statistics: The Science of Uncertainty*. New York: W. H. Freeman and Company.
Field, Hartry (1978), "A Note on Jeffrey Conditionalization," *Philosophy of Science* 45: 361-367.
Genest, Christian, Kevin J. McConway, and Mark J. Schervish (1986), "Characterization of Externally Bayesian Pooling Operators," *The Annals of Statistics* 14: 487-501.
Good, Irving J. (1950), *Probability and the Weighing of Evidence*. London: Charles Griffin.
______(1983), *Good Thinking*, Minneapolis: University of Minnesota Press.
Iverson, Kenneth E. (1962), *A Programming Language*, New York: Wiley.
Jeffrey, Richard (1992a), *Probability and the Art of Judgment*, Cambridge: Cambridge University Press.
______(1992b), "Postscript 1991: New Explanation Revisited," in Jeffrey 1992a, 103-107.
______(1995), Probability Reparation: The Problem of New Explanation," *Philosophical Studies* 77: 97-102.
______(2004), *Subjective Probability: The Real Thing*, Cambridge: Cambridge University Press.

Joyce, James (2004), "On the Plurality of Probabilist Measures of Confirmation," unpublished manuscript presented at the PSA 2004 symposium on inductive logic.
Knuth, Donald E. (1992), "Two Notes on Notation," *The American Mathematical Monthly* 99: 403-422.
Lange, Marc (2000), "Is Jeffrey Conditionalization Defective in Virtue of Being Non-Commutative? Remarks on the Sameness of Sensory Experience," *Synthese* 123: 393-403.
Madansky, Albert (1964), "Externally Bayesian Groups," Technical Report RM-4141-PR, RAND Corporation.
______(1978), "Externally Bayesian Groups," Unpublished manuscript, University of Chicago.
Osherson, Daniel (2002), "Order Dependence and Jeffrey Conditionalization," unpublished paper available at: http://www.princeton.edu/~osherson/papers/jeff3.pdf .
Raiffa, Howard (1968), *Decision Analysis: Introductory Lectures on Choices under Uncertainty.* Reading, MA: Addison-Wesley.
Royden, H. L. (1963), *Real Analysis*, New York: The Macmillan

Company.
Wagner, Carl G. (1997), "Old Evidence and New Explanation," *Philosophy of Science* 64: 677-691.
______(1999), "Old Evidence and New Explanation II," *Philosophy of Science* 66: 283-288.
______(2001), "Old Evidence and New Explanation III," Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association, *Philosophy of Science* 68: S165-S175.
______(2002), "Probability Kinematics and Commutativity," *Philosophy of Science* 69: 266-278.
______(2003), "Commuting Probability Revisions: The Uniformity Rule," *Erkenntnis* 59: 349-364.