*Philosophy of Science* (2022), **89**, 631–643
doi:10.1017/psa.2022.1

DISCUSSION NOTE

# Postscript to Richard Jeffrey's "Conditioning, Kinematics, and Exchangeability"

Carl G. Wagner

Department of Mathematics, University of Tennessee, 1403 Circle Drive, Ayres Hall 227, Knoxville, TN, US. Email: cwagner@tennessee.edu

(Received 03 March 2021; revised 04 June 2021; accepted 07 July 2021; first published online 11 February 2022)

## Abstract

Richard Jeffrey's "Conditioning, Kinematics, and Exchangeability" is one of the foundational documents of probability kinematics. However, the section entitled "Successive Updating" contains a subtle error regarding the applicability of updating by so-called *relevance quotients* in order to ensure the commutativity of successive probability kinematical revisions. Upon becoming aware of this error, Jeffrey formulated the appropriate remedy, but he never discussed the issue in print. To head off any confusion, it seems worthwhile to alert readers of Jeffrey's "Conditioning, Kinematics, and Exchangeability" to the aforementioned error and to document his remedy, placing it in the context of both earlier and subsequent work on commuting probability kinematical revisions.[1]

## 1. Introduction

Along with Richard Jeffrey's book *The Logic of Decision* (1983) and the mathematically bountiful article "Updating Subjective Probability" (Diaconis and Zabell 1982), Jeffrey's "Conditioning, Kinematics, and Exchangeability" (1988, 1992) is one of the foundational documents of probability kinematics. Among other things, it gives a beautifully lucid account of various equivalent formulations of the preconditions for updating a prior by conditioning or by probability kinematics. However, the section entitled "Successive Updating" contains a subtle error regarding the applicability of updating by so-called *relevance quotients* in order to ensure the commutativity of successive probability kinematical revisions. Upon becoming aware of this error, Jeffrey formulated the appropriate remedy, but he never discussed the issue in print. To head off any confusion, it seems worthwhile to alert readers of Jeffrey's "Conditioning, Kinematics, and Exchangeability" to the aforementioned error and to document his remedy, placing it in the context of both earlier and subsequent work on commuting probability kinematical revisions.

Although this Discussion Note touches on some of the philosophical and methodological issues that arise in choosing the correct representation of what is learned

[1] This Discussion Note summarizes and elaborates on discussions between Jeffrey and the author that occurred during the late 1990s. The core of those discussions is described in section 5.

from new evidence alone, it is intended primarily to clarify certain mathematical aspects of successive probability kinematical revisions. More detailed discussions of associated philosophical issues may be found in Field (1978), Lange (2000), Wagner (2002), and Hawthorne (2004).

## 2. Notation and terminology

In what follows, $\Omega$ denotes a set of possible states of the world, conceived as mutually exclusive and exhaustive, and **A** denotes an algebra of subsets (called *events*) of $\Omega$. If $p$ and $q$ are finitely additive probability measures on **A** and $A \in \mathbf{A}$, the *relevance quotient* (terminology attributed to Carnap), denoted by $R_p^q(A)$, is defined by the formula $R_p^q(A) := q(A)/p(A)$. Typically, $q$ is thought of as resulting from the revision of $p$ as a result of encountering new evidence. In such cases, $p$ is called the *prior probability*, and $q$ is called the *posterior probability*. If $q$ comes from $p$ by conditioning on the event $E$, then

$$R_p^q(A) = p(A|E)/p(A) = p(A \cap E)/p(A)p(E) = p(E|A)/p(E).$$

Note that $R_p^q(A)$ contains implicit restraints on the prior $p$. As a simple example, if $R_p^q(A) = 2$, then, necessarily, $p(A) \leq 1/2$. We will return to this apparently trivial observation later in this note.

If $p$ and $q$ are as stated previously, and $A$ and $B$ are events, the *Bayes factor*, denoted by $B_p^q(A:B)$, is defined by the formula $B_p^q(A:B) := \frac{q(A)/q(B)}{p(A)/p(B)}$, which is simply the ratio of the new to old odds on $A$ against $B$. Relevance quotients and Bayes factors are connected by the formula

$$B_p^q(A:B) = \frac{R_p^q(A)}{R_p^q(B)}. \tag{1}$$

When $q$ comes from $p$ by conditioning on $E$, then $B_p^q(A:B)$ reduces to the familiar *likelihood ratio* $p(E|A)/p(E|B)$.

## 3. Probability kinematics

In the remainder of this note, all probability measures are assumed to be *strictly coherent*, in the sense that every nonempty event $A$ is assigned a nonzero probability. This assumption, although inessential, allows us to avoid the distraction of continually having to postulate the positivity of various probabilities in theorems and their proofs.

Suppose that $p$ is your prior probability on **A**, and $\mathbf{E} = \{E_1, \ldots, E_n\}$ is a partition of $\Omega$, with each $E_i \in \mathbf{A}$. New evidence prompts you to revise $p$ to the posterior probability measure $q$ as follows. Based on the total evidence, old as well as new, you first assess the posterior probabilities $q(E_i) = e_i$, where, of course, $e_1 + \cdots + e_n = 1$. Judging that you have learned nothing that would disturb any of the prior conditional probabilities $p(A|E_i)$, you adopt the *rigidity condition* $q(A|E_i) = p(A|E_i)$, for all $A \in \mathbf{A}$ and $i = 1, \ldots, n$. This fully and uniquely determines $q$ (Jeffrey 1983) by the formula

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

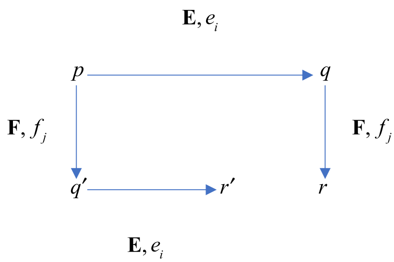


**Figure 1.** Elementary successive updating.

$$q(A) = \sum_{i=1}^{n} e_i p(A|E_i). \tag{2}$$

When probability measures $q$ and $p$ are related by equation (2), we say that $q$ has come from $p$ by *probability kinematics* (henceforth, PK), or by *Jeffrey conditioning*, on the partition $\mathbf{E}$.

## 4. Successive updating

### 4.1. *The elementary model*

Consider two possible successive updating schemes. In the first instance, $p$ is first revised to $q$ by PK on the partition $\mathbf{E} = \{E_1, \ldots, E_n\}$ of $\Omega$, with $q(E_i) = e_i$, and then $q$ is revised to $r$ on the partition $\mathbf{F} = \{F_1, \ldots, F_m\}$ of $\Omega$, with $r(F_j) = f_j$. In the second instance, $p$ is first revised to $q'$ by PK on $\mathbf{F}$, with $q'(F_j) = f_j$, and then $q'$ is revised to $r'$ by PK on $\mathbf{E}$, with $r'(E_i) = e_i$ (figure 1).

If it turns out that $r' = r$, the successive PK revisions are said to *commute*. It is straightforward to verify that

$$r(A) = \sum_{i,j} \frac{e_i f_j}{p(E_i)q(F_J)} p(AE_iF_j) \tag{3}$$

$$r'(A) = \sum_{i,j} \frac{e_i f_j}{q'(E_i)p(F_J)} p(AE_iF_j). \tag{4}$$

It is obvious from equations (3) and (4) that the conditions $q'(E_i) = p(E_i)$ and $q(F_j) = p(F_j)$, which Diaconis and Zabell (1982) dub with the beautifully suggestive nomenclature *Jeffrey independence*, are sufficient to ensure commutativity. In fact, they prove that Jeffrey independence is necessary for commutativity as well. In general, however, $r'$ may differ from $r$. Individuals who have found this troubling (see Lange [2000] for some sample references) presumably subscribe to the following two principles:

1. If the revisions of $p$ to $q$, and of $q'$ to $r'$, are based on identical new learning, and the revisions of $q$ to $r$, and of $p$ to $q'$, are based on identical new learning, then it ought to be the case that $r' = r$.

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

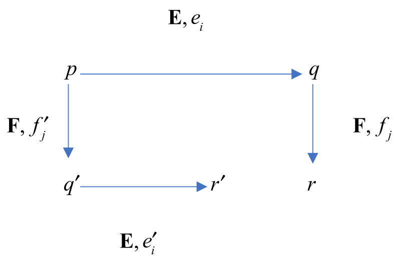


**Figure 2.** Extended successive updating.

2. Identical new learning prompting the revisions of $p$ to $q$, and of $q'$ to $r'$, should be represented by the identities $r'(E_i) = q(E_i) = e_i$, for all $i$. Identical new learning prompting the revisions of $q$ to $r$, and of $p$ to $q'$, should be represented by the identities $q'(F_j) = r(F_j) = f_j$, for all $j$.

Although the first of these principles seems uncontroversial, the second is profoundly mistaken. This was already noted by Carnap in correspondence with Jeffrey in the late 1950s, as described by Jeffrey (1975). Carnap pointed out (in the terminology of our current example) that the probabilities $r'(E_i)$ are based not only on the new learning prompting the revision of the probabilities $q'(E_i)$ but also on the totality of evidence incorporated in the latter probabilities. Similar remarks apply, of course, to the probabilities $q'(F_j)$. Carnap's point was forcefully reiterated by Field (1978), who proceeded to identify the correct representation of what is learned from new evidence alone, the details of which we examine in the next subsection.

### 4.2. *The extended model: Field's analysis*

The term *extended model* refers to the generalization of figure 1 shown in figure 2. As our notation suggests, it is no longer assumed in the extended model that $e'_i = e_i$ or that $f'_j = f_j$. Under what conditions do we get commutativity in this model? Hartry Field (1978), presumably inspired by the old Bayesian idea (Good 1950, 1983) that ratios of new to old *odds* furnish the correct representation of what is learned from new evidence alone, established the remarkable result that the classical PK formula in equation (2) could be transformed into a "re-parameterized" form:

$$q(A) = \sum_{i=1}^{n} G_i p(AE_i) \Big/ \sum_{i=1}^{n} G_i p(E_i), \text{ where } G_i := \left( \prod_{k=1}^{n} B_p^q(E_i : E_k) \right)^{1/n}. \quad (5)$$

[2]

[2] Note that $G_i$ is simply the geometric mean of certain Bayes factors. Field actually expressed $G_i$ in the form $e^{\alpha_i}$, where $\alpha_i = \ln G_i$, and interpreted $\alpha_i$ as expressing the direct and immediate effect of a given stimulus. However, Garber (1980) noted that if $\alpha_i > 0$, repeated exposure to that stimulus would then drive the value of $q(E_i)$ toward 1. Unfortunately, Garber's counterexample led philosophers to ignore Field's beautiful re-parameterization of JC, which, as we'll see, clearly, and for the first time, exhibited sufficient conditions for successive JC revisions on different partitions to commute. As argued by Wagner (2002), Field's analysis can easily be divested of its physicalist gloss.

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

The analogous re-parameterizations of the classical formulas for $r, q'$, and $r'$ are

$$r(A) = \sum_{j=1}^{m} g_j q(AF_j) / \sum_{j=1}^{m} g_j q(F_j), \text{ where } g_j := \left( \prod_{k=1}^{m} B_q^r(F_j : F_k) \right)^{1/m} \tag{6}$$

$$q'(A) = \sum_{j=1}^{m} g'_j p(AF_j) / \sum_{j=1}^{m} g'_j p(F_j), \text{ where } g'_j := \left( \prod_{k=1}^{m} B_p^{q'}(F_j : F_k) \right)^{1/m}, \tag{7}$$

and

$$r'(A) = \sum_{i=1}^{n} G'_i q'(AE_i) / \sum_{i=1}^{n} G'_i q'(E_i) \text{ where } G'_i := \left( \prod_{k=1}^{n} B_{q'}^{r'}(E_i : E_k) \right)^{1/n}. \tag{8}$$

Combining equations (5)–(8) yields the successive PK revision formulas

$$r(A) = \sum_{i,j} G_i g_j p(AE_i F_j) / \sum_{i,j} G_i g_j p(E_i F_j) \tag{9}$$

and

$$r'(A) = \sum_{i,j} G'_i g'_j p(AE_i F_j) / \sum_{i,j} G'_i g'_j p(E_i F_j), \tag{10}$$

from which the following theorem follows immediately:

**Theorem 1.** The *Field parameter identities*

$$G'_i = G_i, \text{ for } 1 \leq i \leq n, \text{ and } g'_j = g_j, \text{ for } 1 \leq j \leq m, \tag{11}$$

imply that $r' = r$.

**Proof.** Obvious. □

It is important not to read more into these results than has so far been established. In figure 2, it is assumed that $p, q, r, q'$, and $r'$ are *fully defined* probability measures on the algebra **A**; that $q$ has come from $p$, and $r'$ from $q'$, by PK on **E**; and that $r$ has come from $q$, and $q'$ from $p$, by PK on **F**. Then, if it is determined that $G'_i = G_i$ and $g'_j = g_j$, it follows that $r' = r$.

Consider, however, the following different scenario, in which the preceding assumptions hold only for $p, q$, and $r$, and the parameters $G_i$ and $g_j$ have been determined. Can we then *design* revisions $q'$ of $p$ by PK on **F**, and $r'$ of $q'$ by PK on **E**, so that we are guaranteed to have $r' = r$? The natural move is to define $q'$ and $r'$ by the formulas

$$q'(A) := \sum_{j=1}^{m} g_j p(AF_j) / \sum_{j=1}^{m} g_j p(F_j) \tag{12}$$

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

and

$$r'(A) := \sum_{i=1}^{n} G_i q'(AE_i) / \sum_{i=1}^{n} G_i q'(E_i). \tag{13}$$

To show that this is the right move, however, requires a proof of the following theorems. First, we need to take note of a key property of products of Field parameters.

**Theorem 2.** $\prod_{i=1}^{n} G_i = \prod_{i=1}^{n} G'_i = \prod_{j=1}^{m} g_j = \prod_{j=1}^{m} g'_j = 1.$

**Proof.** By equations (5) and (1), $\prod_{i=1}^{n} G_i = \prod_{i=1}^{n} \left( \prod_{k=1}^{n} R_p^q(E_i)/R_p^q(E_k) \right)^{1/n} = \left( \prod_{i=1}^{n} R_p^q(E_i)^n / (\prod_{k=1}^{n} R_p^q(E_k))^n \right)^{1/n} = 1^{1/n} = 1.$ □

Next, we can show that the probabilities defined by equations (12) and (13) behave just as we intend.

**Theorem 3.** (i) The set function $q'$ defined by equation (12) is a probability measure on **A** and comes from $p$ by PK on **F**. Moreover,

$$g'_j := \left( \prod_{k=1}^{m} B_p^{q'}(F_j : F_k) \right)^{1/m} = g_j. \tag{14}$$

(ii) The set function $r'$ defined by equation (13) is a probability measure on **A**, and $r'$ comes from $q'$ by PK on **E**. Moreover,

$$G'_i := \left( \prod_{k=1}^{n} B_{q'}^{r'}(E_i : E_k) \right)^{1/n} = G_i. \tag{15}$$

So by theorem 2, $r' = r$.

**Proof.** (i) It is easy to show that $q'$ is an additive set function on **A** and that $q'(\Omega) = 1$. Also,
$q'(A|F_j) = q'(AF_j)/q'(F_j) = [g_j p(AF_j)/\sum_{j=1}^{m} g_j p(F_j)]/[g_j p(F_j)/\sum_{j=1}^{m} g_j p(F_j)] = p(A|F_j),$
and so $q'$ comes from $p$ by PK on **F**. Finally, equation (13) implies that $B_p^{q'}(F_j : F_k) = g_j/g_k$, and so $\left( \prod_{k=1}^{m} B_p^{q'}(F_j : F_k) \right)^{1/m} = (g_j^m / g_1 \cdots g_m)^{1/m} = g_j$, by theorem 1. The proof of (ii) is similar. □

In the next section, we will encounter an attempt to simplify Field's analysis for which analogues of equations (14) and (15) fail to obtain.

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

## 5. Jeffrey's proposal

In an attempt to simplify Field's parameterization of JC, Jeffrey noted that in the extended model, the classical PK formula $q(A) = \sum_{i=1}^{n} e_i p(A|E_i)$ can be recast as $q(A) = \sum_{i=1}^{n} R_i p(AE_i)$, where $R_i = R_p^q(E_i)$. Similarly, one can recast the classical formulas for $r, q'$, and $r'$ as $r(A) = \sum_{j=1}^{m} \rho_j q(AF_j)$, $q'(A) = \sum_{j=1}^{m} \rho'_j p(AF_j)$, and $r'(A) = \sum_{i=1}^{n} R'_i q'(AE_i)$, where $\rho_j = R_q^r(F_j)$, $\rho'_j = R_p^{q'}(F_j)$, and $R'_i = R_{q'}^{r'}(E_i)$. It follows that

$$r(A) = \sum_{i,j} R_i \rho_j p(AE_iF_j), \text{ and } r'(A) = \sum_{i,j} R'_i \rho'_j p(AE_iF_j). \tag{16}$$

So if the *relevance quotient identities*

$$R'_i = R_i,\ i = 1, \ldots, n \text{ and } \rho'_j = \rho_j, j = 1, \ldots, m \tag{17}$$

hold, then $r' = r$.

Again, it is important to keep in mind here that this commutativity result depends on the assumption that $p$, $q$, $r$, $r$, $q'$, and $r'$ are *fully defined* probability measures on the algebra **A**; that $q$ has come from $p$, and $r'$ from $q'$, by PK on **E**; and that $r$ has come from $q$, and $q'$ from $p$, by PK on **F**. Then, if it is determined that $R'_i = R_i$ and $\rho'_j = \rho_j$, it follows that $r' = r$. But suppose that only $p$, $q$, and $r$ have been assessed and the relevance quotients $R_i$ and $\rho_j$ have been evaluated. Can we then *design* PK revisions of $p$ to $q'$on **F**, and of $q'$to $r'$on **E**, so that we are guaranteed to have $r' = r$? Jeffrey proposed setting $q'(F_j)$ equal to $\rho_j p(F_j)$ and setting $r'(E_i)$ equal to $R_i q'(E_i)$. That this may sometimes fail to do the trick can be seen from the example in Jeffrey's table 2 (1988, 236; 1992, 134). In this example, $\Omega = \{1, 2, 3, 4\}, E_1 = \{1, 2\}, E_2 = \{3, 4\}$, $F_1 = \{1, 4\}$, and $F_2 = \{2, 3\}$. The prior $p$ is defined by $p(i) = i/10$, for $i = 1, \ldots, 4$. The probability measure $q$ comes from $p$ by PK on $\{E_1, E_2\}$, with $q(E_1) = q(E_2) = 1/2$, and the probability measure $r$ comes from $q$ by PK on $\{F_1, F_2\}$, with $r(F_1) = r(F_2) = 1/2$. In table 1, the distracting arithmetic mistakes in Jeffrey's table have been corrected (with corrected values in parentheses) so that the error in his proposal for defining $q'(F_j)$ and $r'(E_i)$ stands out more clearly.

Note that we do in fact arrive at $r' = r$. This was, of course, predictable, in view of the commutativity of ordinary multiplication. But an odd thing occurs on the path from $p$ to $r'$: we pass through what we have labeled "$q'$," which fails to define a probability measure because its entries do not sum to 1. This is simply an illustration of the fact, remarked upon in section 2, that the relevance quotient $R_p^q(A)$ contains implicit constraints on the prior probability $p(A)$. So although the positive real numbers $\rho_1, \ldots, \rho_n$ *might* function as a sequence of relevance quotients, in the sense that there exist probabilities $\pi_1, \ldots, \pi_n$ with $\pi_1 + \cdots + \pi_n = 1$ and$\rho_1\pi_1 + \cdots + \rho_n\pi_n = 1$, this need not be the case for *every* sequence of probabilities that sum to 1, just as we saw in table 1.

Upon becoming aware of this problem, Jeffrey proposed to repair the array marked "$q'$" by normalizing, that is, by dividing each of its entries by 441/437, which then defines, by the array marked $q'$ in table 2, a genuine probability measure. But

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

**Table 1.** Jeffrey's Table 2 (arithmetic corrected)

$p$

| | $F_1$ | $F_2$ | | |
|---|---|---|---|---|
| $E_1$ | $\frac{1}{10}$ | $\frac{2}{10}$ | $\frac{3}{10}$ | $R_1 = \frac{5}{3}$ |
| $E_2$ | $\frac{4}{10}$ | $\frac{3}{10}$ | $\frac{7}{10}$ | $R_2 = \frac{5}{7}$ |
| | $\frac{1}{2}$ | $\frac{1}{2}$ | $1$ | |

$q$

| | $F_1$ | $F_2$ | |
|---|---|---|---|
| $E_1$ | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{2}$ |
| $E_2$ | $\frac{2}{7}$ | $\frac{3}{14}$ | $\frac{1}{2}$ |
| | $\frac{19}{42}$ | $\frac{23}{42}$ | $1$ |
| | $\rho_1 = \frac{21}{19}$ | $\rho_2 = \left(\frac{21}{23}\right)$ | |

"$q'$"

| | $F_1$ | $F_2$ | |
|---|---|---|---|
| $E_1$ | $\frac{21}{190}$ | $\left(\frac{21}{115}\right)$ | $\left(\frac{1281}{4370}\right)$ |
| $E_2$ | $\frac{42}{95}$ | $\left(\frac{63}{230}\right)$ | $\left(\frac{3129}{4370}\right)$ |
| | $\frac{21}{38}$ | $\left(\frac{21}{46}\right)$ | $\left(\frac{441}{437}!\right)$ |

$r' = r$

| | $F_1$ | $F_2$ | |
|---|---|---|---|
| $E_1$ | $\frac{7}{38}$ | $\left(\frac{7}{23}\right)$ | $\left(\frac{427}{874}\right)$ |
| $E_2$ | $\frac{6}{19}$ | $\left(\frac{9}{46}\right)$ | $\left(\frac{447}{874}\right)$ |
| | $\frac{1}{2}$ | $\frac{1}{2}$ | $(1)$ |

now, if the entries in the first row of $q'$ are multiplied by 5/3, and the entries in the second row are multiplied by 5/7, the resulting array fails to define a probability measure because its entries, predictably, sum to 437/441. Dividing every entry in that table by 437/441 then defines, by the array marked $r'$ in table 2, a genuine probability measure. Moreover, $r' = r$, as intended.

Notice that the commutativity in table 2 is, contrary to what Jeffrey had hoped for,[3] no longer accounted for by *relevance quotient identities*. What we get instead are the *relevance quotient proportionalities* $\rho'_j \propto \rho_j$ and $R'_i \propto R_i$, with

$$\rho'_j = (437/441) \cdot \rho_j, \; j = 1, 2 \quad \text{and} \quad R'_i = (441/437) \cdot R_i, \; i = 1, 2. \tag{18}$$

As we will see in section 7, analogous proportionalities prove to be the rule, rather than the exception, in the most general parameterization of JC.

[3] "Updating is always commutative when taking a step is a matter of *setting ratios* ... of new to old cell probabilities" (Jeffrey 1988, 236; 1992, 134).

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

**Table 2.** Table 1, rectified by normalizing

$$p = \begin{array}{c|ccc} & F_1 & F_2 & \\ E_1 & \frac{1}{10} & \frac{2}{10} & \frac{3}{10} \\ E_2 & \frac{4}{10} & \frac{3}{10} & \frac{7}{10} \\ & \frac{1}{2} & \frac{1}{2} & 1 \end{array} \qquad R_1 = \frac{5}{3} \quad R_2 = \frac{5}{7} \qquad \longrightarrow \qquad \begin{array}{c|ccc} & F_1 & F_2 & \\ E_1 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ E_2 & \frac{2}{7} & \frac{3}{14} & \frac{1}{2} \\ & \frac{19}{42} & \frac{23}{42} & 1 \end{array} = q$$

$$\rho'_1 = \frac{23}{21} \quad \rho'_2 = \frac{19}{21} \ \Big\downarrow \qquad\qquad \rho_1 = \frac{21}{19} \quad \rho_2 = \frac{21}{23} \ \Big\downarrow$$

$$q' = \begin{array}{c|ccc} & F_1 & F_2 & \\ E_1 & \frac{23}{210} & \frac{437}{2415} & \frac{61}{210} \\ E_2 & \frac{46}{105} & \frac{19}{70} & \frac{149}{210} \\ & \frac{23}{42} & \frac{19}{42} & 1 \end{array} \qquad R'_1 = \frac{735}{437} \quad R'_2 = \frac{315}{437} \qquad \longrightarrow \qquad r' = \begin{array}{c|ccc} & F_1 & F_2 & \\ E_1 & \frac{7}{38} & \frac{7}{23} & \frac{427}{874} \\ E_2 & \frac{6}{19} & \frac{9}{46} & \frac{447}{874} \\ & \frac{1}{2} & \frac{1}{2} & 1 \end{array} = r$$

## 6. The Jeffrey–Hendrickson parameterization of JC

It is ironic that while Jeffrey sought to simplify Field's analysis of commutativity by employing relevance quotients, he had in hand, in Jeffrey and Hendrickson (1988/1989), the perfect parameterization of JC for accomplishing that task. The Jeffrey–Hendrickson transformation of the classical formula in equation (2) takes the form

$$q(A) = \sum_{i=1}^{n} B_i p(AE_i) / \sum_{i=1}^{n} B_i p(E_i), \text{ where } B_i := B_p^q(E_i : E_1). \tag{19}$$

$$r(A) = \sum_{j=1}^{m} b_j q(AF_j) / \sum_{j=1}^{m} b_j q(F_j), \text{ where } b_j := B_q^r(F_j : F_1), \tag{20}$$

$$q'(A) = \sum_{j=1}^{m} b'_j p(AF_j) / \sum_{j=1}^{m} b'_j p(F_j), \text{ where } b'_j := B_p^{q'}(F_j : F_1), \text{ and} \tag{21}$$

$$r'(A) = \sum_{i=1}^{n} B'_i q'(AE_i) / \sum_{i=1}^{n} B'_i q'(E_i), \text{ where } B'_i := B_{q'}^{r'}(E_i : E_1), \tag{22}$$

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

from which it follows that

$$r'(A) = \sum_{i,j} B'_i b'_j p(AE_iF_j) / \sum_{i,j} B'_i b'_j p(E_iF_j), \text{ and} \tag{23}$$

$$r(A) = \sum_{i,j} B_i b_j p(AE_iF_j) / \sum_{i,j} B_i b_j p(E_iF_j). \tag{24}$$

**Theorem 4.** The *Jeffrey–Hendrickson parameter identities*

$$B'_i = B_i, \text{ for } 1 \le i \le n, \text{ and } b'_j = b_j, \text{ for } 1 \le j \le m, \tag{25}$$

are sufficient and, under the regularity conditions,

$$\forall i_1 \forall i_2 \exists j : p(E_{i_1}F_j)p(E_{i_2}F_j) > 0, \text{ and} \tag{26}$$

$$\forall j_1 \forall j_j \exists i : p(E_iF_{j_1})p(E_iF_{j_2}) > 0, \tag{27}$$

necessary for $r' = r$.

**Proof.** See Wagner (2002, theorems 3.1 and 4.1). □

Here again, commutativity depends on the assumption that $p$, $q$, $r$, $q'$, and $r'$ are *fully defined* probability measures on the algebra **A**; that $q$ has come from $p$, and $r'$ from $q'$, by PK on **E**; and that $r$ has come from $q$, and $q'$ from $p$, by PK on **F**. Suppose, however, that only $p$, $q$, and $r$ have been defined, and the parameters $B_i$ and $b_j$ have been determined. As in the case of Field's parameterization, we can then *design* probability measures $q'$ and $r'$ by means of the definitions

$$q'(A) := \sum_{j=1}^{m} b_j p(AF_j) / \sum_{j=1}^{m} b_j p(F_j) \tag{28}$$

and

$$r'(A) := \sum_{i=1}^{n} B_i q'(AE_i) / \sum_{i=1}^{n} B_i q'(E_i) \tag{29}$$

so that the following analogue of theorem 3 holds:

**Theorem 5.** (i) The set function $q'$ defined by equation (28) is a probability measure on **A** and comes from $p$ by PK on **F**. Moreover, $b'_j := B_p^{q'}(F_j : F_1) = b_j$. (ii) The set function $r'$ defined by equation (29) is a probability measure on **A**, and $r'$ comes from $q'$ by PK on **E**. Moreover, $B'_i := B_{q'}^{r'}(E_i : E_1) = B_i$. Hence, (iii) $r' = r$.

**Proof.** The proofs of (i) and (ii) are straightforward, and (iii) then follows from theorem 4. □

https://doi.org/10.1017/psa.2022.1 Published online by Cambridge University Press

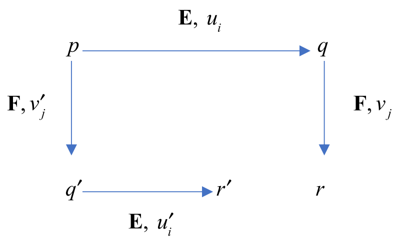


**Figure 3.** Generalized successive PK updating.

## 7. A comprehensive parameterization of JC

Consider the formula $\hat{q}(A) = \sum_{i=1}^{n} u_i p(AE_i)$, where $p$ is a probability measure on A, and the parameters $u_i$ are *any positive real numbers whatsoever*. It is easy to check that $\hat{q}$ is a nonnegative, additive set function on **A**. So $\hat{q}$ is a probability measure if and only if $\hat{q}(\Omega) = \sum_{i=1}^{n} u_i p(E_i) = 1$. Consequently, whatever the value of $\sum_{i=1}^{n} u_i p(E_i)$ turns out to be (whether equal to 1 or not), the set function $q$, defined by

$$q(A) = \hat{q}(A)/\hat{q}(\Omega) = \sum_{i=1}^{n} u_i p(AE_i) / \sum_{i=1}^{n} u_i p(E_i), \tag{30}$$

is a probability measure on **A**. Moreover, because $q(A|E_i) = p(A|E_i)$, for all $A \in \mathbf{A}$ and $1 \leq i \leq n$, $q$ comes from $p$ by PK on **E**.

Suppose now that **E**$= \{E_1, \ldots, E_n\}$, **F**$= \{F_1, \ldots, F_m\}$, and $(u_i)_{1 \leq i \leq n}$, $(u'_i)_{1 \leq i \leq n}$, $(v_j)_{1 \leq j \leq m}$, and $(v'_j)_{1 \leq j \leq m}$ are sequences of arbitrary positive real numbers. Consider the successive PK updating scenario shown in figure 3.

In figure 3, the probability measure $q$ comes from $p$ by PK on **E** in accord with the formula in equation (30). Similarly, $r$ comes from $q$ by PK on **F**, $q'$ comes from $p$ by PK on **F**, and $r'$comes from $q'$ by PK on **E** by the analogous formulas

$$r(A) = \sum_{j=1}^{m} v_j q(AF_j) / \sum_{j=1}^{m} v_j q(F_j), \tag{31}$$

$$q'(A) = \sum_{j=1}^{m} v'_j p(AF_j) / \sum_{j=1}^{m} v'_j p(F_j), \tag{32}$$

and

$$r'(A) = \sum_{i=1}^{n} u'_i q'(AE_i) / \sum_{i=1}^{n} u'_i q'(E_i). \tag{33}$$

It follows that

$$r(A) = \sum_{i,j} u_i v_j p(AE_iF_j) / \sum_{i,j} u_i v_j p(E_iF_j), \text{ and} \tag{34}$$

$$r'(A) = \sum_{i,j} u_i' v_j' p(AE_iF_j) / \sum_{i,j} u_i' v_j' p(E_iF_j). \tag{35}$$

From equations (34) and (35), a condition sufficient to ensure that $r' = r$ is obvious.

**Theorem 6.** If there exists a constant $c$ such that $u_i' = c \cdot u_i$, for $1 \leq i \leq n$ (symbolized by $u_i' \propto u_i$), and there exists a constant $d$ such that $v_j' = d \cdot v_j$ for $1 \leq j \leq m$ (symbolized by $v_j' \propto v_j$), then $r' = r$.

**Proof.** Straightforward. □

The proportionalities $u_i' \propto u_i$ and $v_j' \propto v_j$ turn out to be equivalent to certain Bayes factor identities. In order to prove this assertion, however, we need to establish a few preliminary results. We begin by establishing a connnection between the rather abstract quantities $u_i$ appearing in the formula in equation (30) and certain Bayes factors.

**Theorem 7.** For all $1 \leq i, k \leq n$, $u_i/u_k = B_p^q(E_i : E_k)$.

**Proof.** By the definition of $B_p^q(E_i : E_k)$, along with the formula in equation (30), we have

$$B_p^q(E_i : E_k) = \frac{q(E_i)p(E_k)}{q(E_k)p(E_i)} = \frac{u_i p(E_i) / \sum_{i=1}^{n} u_i p(E_i)}{u_k p(E_k) / \sum_{i=1}^{n} u_i p(E_i)} \times \frac{p(E_k)}{p(E_i)} = \frac{u_i}{u_k}.$$ □

**Remark.** Analogous formulas for $v_j/v_k$, as well as for $u_i'/u_k'$and $v_j'/v_k'$, should be obvious.

**Theorem 8.** In the successive updating scenario displayed in figure 3, the proportionality $u_i' \propto u_i$ is equivalent to the *Bayes factor identities*

$$B_{q'}^{r'}(E_i : E_k) = B_p^q(E_i : E_k), \text{ for } 1 \leq i, k \leq n, \tag{36}$$

and the proportionality $v_j' \propto v_j$ is equivalent to the Bayes factor identities

$$B_p^{q'}(F_j : F_k) = B_q^r(F_j : F_k), \text{ for } 1 \leq j, k \leq m. \tag{37}$$

**Proof.** Suppose first that $u_i' \propto u_i$, so that there exists a constant $c$ such that $u_i' = c \cdot u_i$, for $i = 1, \ldots, n$. By theorem 7, $B_{q'}^{r'}(E_i : E_k) = \frac{u_i'}{u_k'} = \frac{c \cdot u_i}{c \cdot u_k} = \frac{u_i}{u_k} = B_p^q(E_i : E_k)$. By theorem 7, equation (36), with $k = 1$, yields $\frac{u_i'}{u_1'} = \frac{u_i}{u_1}$, whence $u_i' = c \cdot u_i$, where $c = u_1'/u_1$. The proof that $v_j' \propto v_j$ is equivalent to equation (37) is nearly identical. □

The probability kinematical formulas in equations (30)–(33) encompass, inter alia, (i) Field's parameterizations ($u_i = G_i$, $v_j = g_j$, etc.); (ii) Jeffrey's parameterizations, after normalization ($u_i = R_i$, $v_j = \rho_j$, etc.); and (iii) the Jeffrey–Hendrickson parameterizations ($u_i = B_i$, $v_j = b_j$, etc.).

In all of these cases, we have exhibited conditions sufficient to ensure commutativity. But the conditions necessary for commutativity have only been stated for the Jeffrey–Hendrickson parameters. Bayes factor identities play a crucial role in formulating such conditions for other parameterizations, as follows:

1. Recall that under the regularity conditions in equations (26) and (27), commutativity implies the Jeffrey–Hendrickson parameter identities in equation (25).
2. Observe that the identities in equation (25) imply the Bayes factor identities in equations (36) and (37) because $B_p^q(E_i : E_k) = B_i/B_k$, and so forth.
3. Observe that the Bayes factor identities imply the Field parameter identities in equation (11).
4. Recall that, by theorem 8, the Bayes factor identities imply the parameter proportionalities $u_i' \propto u_i$ and $v_j' \propto v_j$.

We conclude the case for the primacy of Bayes factors in the representation of what is learned from new evidence alone with the final observation that in the elementary model of sequential PK revision represented in figure 1, the Bayes factor identities turn out to be both necessary and sufficient for Jeffrey independence.

## References


Diaconis, Persi, and Sandy Zabell. 1982. "Updating Subjective Probability." *Journal of the American Statistical Association* 77 (380):822–30.

Field, Hartry. 1978. "A Note on Jeffrey Conditionalization." *Philosophy of Science* 45 (3):361–67.

Garber, Daniel. 1980. "Field and Jeffrey Conditionalization." *Philosophy of Science* 47 (1):142–45.

Good, Irving J. 1950. *Probability and the Weighing of Evidence*. London: Charles Griffin.

Good, Irving J. 1983. *Good Thinking*. Minneapolis: University of Minnesota Press.

Hawthorne, James. 2004. "Three Models of Sequential Belief Updating on Uncertain Evidence." *Journal of Philosophical Logic* 33 (1):89–123.

Jeffrey, Richard. 1975. "Carnap's Empiricism." In *Minnesota Studies in the Philosophy of Science*, vol. VI, edited by Grover Maxwell and Robert M. Anderson Jr., 37–49. Minneapolis: University of Minnesota Press.

Jeffrey, Richard. 1983. *The Logic of Decision*. Chicago: University of Chicago Press.

Jeffrey, Richard. 1988. "Conditioning, Kinematics, and Exchangeability." In *Causation, Chance, and Credence*, vol. 1, edited by B. Skyrms and W. Harper, 221–55. Dordrecht: Reidel.

Jeffrey, Richard. 1992. "Conditioning, Kinematics, and Exchangeability." *Probability and the Art of Judgment*. Cambridge: Cambridge University Press.

Jeffrey, Richard, and Michael Hendrickson. 1988/1989. "Probabilizing Pathology." *Proceedings of the Aristotelian Society* 89 (3):211–25.

Lange, Marc. 2000. "Is Jeffrey Conditionalization Defective in Virtue of Being Non-Commutative? Remarks on the Sameness of Sensory Experience." *Synthese* 123 (3):393–403.

Wagner, Carl. 2002. "Probability Kinematics and Commutativity." *Philosophy of Science* 69 (2):266–78.


**Cite this article:** Wagner, Carl G. 2022. "Postscript to Richard Jeffrey's "Conditioning, Kinematics, and Exchangeability"." *Philosophy of Science* 89 (3):631–643. https://doi.org/10.1017/psa.2022.1