First published in *Boston Studies in the Philosophy of Science, Vol. VIII*, (1971) R. Buck & R. S. Cohen (Eds.), pp. 342-374. Dordrecht: D. Reidel Publishing Company.

New Dimensions of Confirmation Theory II: The Structure of Uncertainty

You are, I am sure, just as aware as I am that the operational nodes of a complex problem, the points at which it can be split open to yield nuggets of new insight or achieve lasting advances, often lie in tediously technical details perhaps incomprehensible to all but specialists in the matter and anyways totally lacking in the romance and easy excitement which attract the topic's dilettantes. I would like you to hold fast to this appreciation, for the concerns I shall raise here are technical indeed. They are, however, *seminal* technicalities which, as we seek to fathom them, should fundamentally reshape our comprehension of the logic of confirmation, causal connectedness, and the foundations of statistical inference. In brief, it will be seen that certain powerful everyday intuitions concerning which propositions are confirmationally irrelevant to which others are exceedingly difficult to justify. When the roots of these intuitions are laid bare, an inherent intimacy emerges between the structure of rational uncertainty and presuppositions about nomic order.

I

Throughout what follows I shall make one important working assumption which, though far from apodictic, is not so controversial as to degrade the value of the sharp focus it makes possible. This is that in a system of rational beliefs, degrees of credibility can be expressed by a measure $\Pr(\_\_ \mid \ldots)$—read "the probability (credibility) that __ , given that ..."—which obeys the axioms of the conditional probability calculus under the latter's standard propositional interpretation. For reasons not unlike certain points raised by Hacking (1967), I am far from convinced that this is an entirely appropriate basis on which to deal with all major issues of confirmation theory. Even so, it may well hold under sufficiently ideal conditions, such as the believer's being aware of all logical relations among the propositions in his belief system, and is in any event the only quantitative credibility model which currently enjoys a modicum of provisional philosophic accord. The distinctive problems and part-solutions here developed in its terms should, I think, persist relatively unmodified into whatever more realistic account of credibility relations may eventually supersede it.

The basic theorems of this model are sufficiently familiar to require no review here. I will, however, remind you that for any background information k and propositions p and q,

$$\text{(1)} \qquad \Pr(p \cdot q \mid k) = \Pr(q \mid k) \times \Pr(p \mid q \cdot k)$$

of which a simple corollary is that

$$\text{(2)} \qquad \frac{\Pr(p \mid q \cdot k)}{\Pr(p \mid k)} = \frac{\Pr(p \cdot q \mid k)}{\Pr(p \mid k) \times \Pr(q \mid k)} = \frac{\Pr(q \mid p \cdot k)}{\Pr(q \mid k)}$$

If we introduce the 'confirmation ratio', CR, by the definition

$$\text{(3)} \qquad \mathrm{CR}(p,\ q \mid k) =_{\text{def}} \frac{\Pr(p \mid q \cdot k)}{\Pr(p \mid k)}$$

as a measure of the degree to which p confirms q relative to background k, (2) may be notationally simplified to

$$\text{(2a)} \qquad \mathrm{CR}(p,\ q \mid k) =_{\text{def}} \frac{\Pr(p \cdot q \mid k)}{\Pr(p \mid k) \times \Pr(q \mid k)} = \mathrm{CR}(q,\ p \mid k)$$

which says that under any background information, p confirms q to exactly the same degree that q confirms p. In particular, insomuch as q confirms, is confirmationally independent of (i.e. is indifferent to), or disconfirms p given background k according to whether $\mathrm{CR}(p, q \mid k)$ is respectively greater than, equal to, or less than unity,[1] (2) entails that whether one proposition confirms, is indifferent to, or disconfirms another is symmetric in the two propositions.

A second principle which will soon be needed is

$$\text{(4)} \qquad \frac{\Pr(p \mid q \cdot r \cdot k)}{\Pr(p \mid k)} = \frac{\Pr(p \mid q \cdot k)}{\Pr(p \mid k)} \times \frac{\Pr(p \mid q \cdot r \cdot k)}{\Pr(p \mid q \cdot k)}$$

the sense of which can perhaps best be grasped from its confirmation-ratio equivalent

$$\text{(4a)} \qquad \mathrm{CR}(p,\ q \cdot r \mid k) = \mathrm{CR}(p,\ q \mid k) \times \mathrm{CR}(p,\ r \mid q \cdot k)$$

In fact, to save later distraction we may as well establish now some more complicated CR-relationships which will eventually prove useful. (The reader who is

[1] I trust that it is essentially uncontroversial by now to explicate confirmation/disconfirmation as an increase/decrease in credibility relative to the background information. But if not, the evident appropriateness of this may perhaps be enhanced by noting that since $\Pr(p \mid k) = \Pr(q \mid k) \times \Pr(p \mid q \cdot k) + \Pr({\sim}\, q \mid k) \times \Pr(p \mid {\sim}\, q \cdot k)$, $0 < \Pr(q \mid k) < 1$ implies that $\mathrm{CR}(p,\ q \mid k)$ is greater or less than unity if and only if $\Pr(p \mid q \cdot k)$ is respectively greater or less than $\Pr(p \mid {\sim}\, q \cdot k)$. That is, so long as q is uncertain given k, q confirms p under k by the CR-criterion iff p's credibility is greater given q and k than given ${\sim}\, q$ and k.

willing to take on faith my proofs of Theorems 1 and 2 below, may omit the remainder of this section.) The proofs of these had best be prefaced with a word about extreme probabilities. Many otherwise straightforward conditional-probability theorems require boundary restrictions to fend off degeneracies which sometimes arise from zero probabilities. (E.g., basic principle $\Pr(p \mid k) = \Pr(p \cdot q \mid k) + \Pr(p \cdot \sim q \mid k)$ fails when k is inconsistent, while the otherwise exceptionless equivalence $\Pr(p \mid q) = \Pr(p \cdot q)/\Pr(q)$ is unreliable when $\Pr(q) = 0$ insomuch as its right-hand side is then ill-defined even though its left-hand side may be perfectly determinate.) To minimize conceptual and visual affront, I will list the nonzero probabilities upon which each result here is conditional in a parenthesized addendum wherein 'p' and 'p/q' abbreviate '$\Pr(p)$' and '$\Pr(p \mid k)$'. Thus '(Nonzero: p, p/q)' following a lemma or theorem advises that it presupposes $\Pr(p) > 0$ and $\Pr(p \mid k) > 0$. Some of the listed boundary restrictions are not strictly necessary, but they can be relaxed only with difficulty and we shall have no need for the increased generality this would afford. In a couple of places, the proofs as here given presuppose additional nonzero probabilities which are neither required to be so by these theorems nor are listed among their nonzero boundary conditions. Modification of the proofs to accommodate zero values of these probabilities should be obvious and will be left to the reader.

For principles (5) and (6), let $r_1, \ldots, r_n$ be a set of mutally exclusive and jointly exhaustive propositions, i.e. $\sum_{i=1}^{n} \Pr(r_i) = 1$ and $Pr(r_i \cdot r_j) = 0$ for $i \neq j$. Then,

(5) $$\sum_{i=1}^{n} \Pr(r_i \mid k) \times \mathrm{CR}(p,\ r_i \mid k) = 1. \quad \text{(Nonzero: } p/k\text{)}.$$

(PROOF: Since $\sum_{i=1}^{n} \Pr(p \cdot r_i \mid k) = \Pr(p \mid k)$ while $\Pr(p \cdot r_i \mid k) = \Pr(p \mid k) \times \Pr(r_i \mid k) \times \mathrm{CR}(p,\ r_i \mid k)$, dividing through by $\Pr(p \mid k)$ establishes the theorem.)

(6) $$\mathrm{CR}(p,\ q) = \sum_{i=1}^{n} \Pr(r_i) \times \mathrm{CR}(p,\ r_i) \times \mathrm{CR}(q, r_i) \times \mathrm{CR}(p,\ q \mid r_i) \quad \text{(Nonzero: } p, q\text{)},$$

or equivalently,

(6a) $$\mathrm{CR}(p,\ q) = \sum_{i=1}^{n} \Pr(r_i \mid p) \times \mathrm{CR}(q, r_i) \times \mathrm{CR}(p,\ q \mid r_i) \quad \text{(Nonzero: } p, q\text{)}.$$

(PROOF: $\Pr(p \cdot q \cdot r_i) = \Pr(r_i) \times \Pr(p \cdot q \mid r_i) = \Pr(r_i) \times \Pr(p \mid r_i) \times \Pr(q \mid r_i) \times [\Pr(p \cdot q \mid r_i)/\Pr(p \mid r_i) \times \Pr(q \mid r_i)] = \Pr(r_i) \times \Pr(p) \times \Pr(q) \times \mathrm{CR}(p, r_i) \times \mathrm{CR}(q, r_i) \times \mathrm{CR}(p,\ q \mid r_i)$. Substituting into $\Pr(p \cdot q) = \sum_{i=1}^{n} \Pr(p \cdot q \cdot r_i)$ then yields (6), while

(6a) follows in turn by noting that $\Pr(r_i) \times \mathrm{CR}(p, r_i) = \Pr(p \cdot r_i)/\Pr(p) = \Pr(r_i \mid p)$. Theorems (6) and (6a) also hold, of course, relative to any consistent background information k.)

Equations (5) and (6) deserve interpretive comment. By letting $\{r_i\}$ be the pair $r, \sim r$, (5) may be seen to justify our intuition that p confirms a proposition iff $\sim p$ disconfirms it. (6) is more surprising, for it shows that $\mathrm{CR}(p, q)$ is guaranteed to lie between $\mathrm{CR}(p,\ q \mid r)$ and $\mathrm{CR}(p,\ q \mid \sim r)$ only if r is confirmationally independent either of p or of q. Otherwise, it is possible e.g. for p to confirm q both given r and given $\sim r$, yet to disconfirm it unconditionally.

LEMMA 1: If $\mathrm{CR}(p,\ q \mid k) = 1$, then $\mathrm{CR}(p,\ \sim q \mid k) = 1$. (Nonzero: p/k, $\sim q/k$).

LEMMA 1′: If $\Pr(q \mid k) = 0$, then $\mathrm{CR}(p,\ \sim q \mid k) = 1$. (Nonzero : p/k).

(PROOF: Let $w = \Pr(q \mid k)$, $a = \mathrm{CR}(p,\ q \mid k)$, $b = \mathrm{CR}(p,\ \sim q \mid k)$. Then from (5) we have $wa + (1 - w)b = 1$, under which $a = 1$ entails $b = 1$ so long as $w < 1$. If $w = 0$, then $b = 1$ unconditionally.)

LEMMA 2: If $\mathrm{CR}(p,\ q \mid r) = \mathrm{CR}(p,\ q \mid \sim r) = 1$, then $\mathrm{CR}(p, q) = 1$ only if either $\mathrm{CR}(p, r) = 1$ or $\mathrm{CR}(q, r) = 1$. (Nonzero: $p,\ q,\ r$).

LEMMA 2′: If $\Pr(p \mid r) = 0$ and $\mathrm{CR}(p,\ q \mid \sim r) = 1$, then $\mathrm{CR}(p, q) = 1$ only if $\mathrm{CR}(p, r) = 1$. (Nonzero: $p,\ q,\ r$).

(PROOF: Let $v = \Pr(r \mid p)$, $w = \Pr(r)$, $a = \mathrm{CR}(q, r)$, $b = \mathrm{CR}(q,\ \sim r)$. Then by (6a) under the lemma's assumptions, $\mathrm{CR}(p, q) = 1$ entails $va + (1 - v)b = 1$. But by (5) we also have $wa + (1 - w)b = 1$. Combining these equations and collecting terms yields $(v - w)(a - b) = 0$, which holds only if either $v = w$ or $a = b$. But $v = w$ entails $\mathrm{CR}(p, r) = 1$ when $w > 0$, while $a = b$ when $wa + (1 - w)b = 1$ only if a, i.e. $\mathrm{CR}(q, r)$, equals 1—which establishes Lemma 2. For Lemma 2′, note from (4a) and (6a) that $\Pr(r \mid q) = 0$ entails $\mathrm{CR}(p, q) = \mathrm{CR}(p,\ \sim r) \times \mathrm{CR}(p,\ q \mid \sim r)$, from which Lemma 2′ is then obvious.)

LEMMA 3: If $\mathrm{CR}(p,\ q \mid r) = \mathrm{CR}(p,\ q \mid \sim r) = 1$, then $\mathrm{CR}(p, q) = 1$ only if either $\mathrm{CR}(q, r) = 1$ or $\Pr(p \mid q \cdot r) = \Pr(p \mid q \cdot \sim r)$. (Nonzero: $p,\ q,\ r,\ \sim r,\ p/r,\ \sim q/r$).

LEMMA 3′: If $\Pr(q \cdot r) = 0$ and $\mathrm{CR}(p, q \mid \sim r) = 1$, then $\mathrm{CR}(p, q) = 1$ only if $\Pr(p \mid q \cdot \sim r) = \Pr(p \mid \sim q \cdot r)$. (Nonzero: $p,\ q,\ r,\ \sim r,\ p/r$).

(PROOF: By Lemma 1, given the stipulated nonzero probabilities, $\mathrm{CR}(p, r) = 1$ and $\mathrm{CR}(p,\ q \mid r) = 1$ respectively entail $\mathrm{CR}(p,\ \sim r) = 1$ and $\mathrm{CR}(p,\ \sim q \mid r) = 1$. Hence by (4a), $\mathrm{CR}(p, r) = 1$ and Lemma 3's assumptions jointly entail $\mathrm{CR}(p,\ q \cdot \sim r) = \mathrm{CR}(p,\ \sim r) \times \mathrm{CR}(p,\ q \mid\sim r) = 1$ and $\mathrm{CR}(p,\ \sim q \cdot r) = \mathrm{CR}(p, r) \times \mathrm{CR}(p, \sim q \mid r) = 1$, whence $\Pr(p \mid q \cdot \sim r) = \Pr(p) = \Pr(p \mid \sim q\ \cdot r)$. Reference to Lemma 2 completes the proof of Lemma 3. Lemma 3′ follows similarly from Lemma 2′ by noting that when $\Pr(r) > 0$, $\Pr(q \cdot r) = 0$ entails $\Pr(q \mid r) = 0$ and hence, by Lemma 1′, $\mathrm{CR}(p,\ \sim q \mid r) = 1$.)

LEMMA 4: If $\mathrm{CR}(p,\ r \mid s) = \mathrm{CR}(p,\ r \mid\sim s) = \mathrm{CR}(q,\ s \mid r) = \mathrm{CR}(q,\ s \mid\sim r) = 1$, and p and r jointly entail q while q and s jointly entail p, then both $\mathrm{CR}(p,\ r) = 1$ and $\mathrm{CR}(q,\ s) = 1$ only if either $\mathrm{CR}(r,\ s) = 1$ or $\Pr(p \equiv q \mid r \cdot \sim s) = \Pr(p \equiv q \mid \sim r \cdot s) = 1$. (Nonzero : $p,\ q,\ r,\ s,\ \sim r,\ \sim s,\ p/r,\ q/s,\ \sim s/r,\ \sim r/s$).

LEMMA 4′: If $\Pr(r \cdot s) = 0$, $\mathrm{CR}(p,\ r \mid \sim s) = \mathrm{CR}(q,\ s \mid \sim r) = 1$, and p and r jointly entail q while q and s jointly entail p, then both $\mathrm{CR}(p, r) = 1$ and $\mathrm{CR}(q, s) = 1$ only if $\Pr(p \equiv q \mid r \cdot \sim s) = \Pr(p \equiv q \mid\sim r \cdot s)$. (Nonzero: $p,\ q,\ r,\ s,\ \sim r,\ \sim s,\ p/r,\ q/s$).

(PROOF: By Lemma 3, Lemma 4's CR-assumptions imply that $\mathrm{CR}(p, r) = \mathrm{CR}(q, s) = 1$ only if either $\mathrm{CR}(r, s), = 1$ or both $\Pr(p \mid r \cdot \sim s) = \Pr(p \mid \sim r\ \cdot s)$ and $\Pr(q \mid r \cdot \sim s) = \Pr(q \mid \sim r \cdot s)$. Similarly, since $\Pr(r \cdot s) = 0$ entails $\Pr(r \mid s) = \Pr(s \mid r) = 0$ when $\Pr(r)$ and $\Pr(s)$ are nonzero, Lemma 3′ and the assumptions of Lemma 4′ imply that $\mathrm{CR}(p, r) = \mathrm{CR}(q, s) = 1$ only if both $\Pr(p \mid r \cdot \sim s) = \Pr(p \mid\sim r\ \cdot s)$ and $\Pr(q \mid\ r \cdot\ \sim s) = \Pr(q \mid\sim r \cdot s)$. To complete the proof, let h_1 and h_2 abbreviate $r \cdot\ \sim s$ and $\sim r\ \cdot s$, respectively. Then if $p \cdot r$ entails q while $q \cdot s$ entails p, we have $\Pr(p \cdot \sim q \mid h_1) = \Pr(\sim p \cdot\ q \mid h_2) = 0$ while $\Pr(p \mid h_1) = \Pr(p \cdot q \mid h_1)$ and $\Pr(q \mid h_2) = \Pr(p \cdot q \mid h_2)$. The latter equations, together with the already established $\Pr(p \mid h_1) = \Pr(p \mid h_2)$ and $\Pr(q \mid h_1) = \Pr(q \mid h_2)$, yield $\Pr(p\ \cdot\ q \mid h_1) = \Pr(p \cdot q \mid h_2) + \Pr(p \cdot \sim q \mid h_2)$ and $\Pr(p \cdot q \mid h_1) + \Pr(\sim p \cdot q \mid h_1) = \Pr(p \cdot q \mid h_2)$, which can hold only if $\Pr(\sim p \cdot q \mid h_1) + \Pr(p \cdot \sim q \mid h_2) = 0$, i.e. only if $\Pr(\sim p \cdot q \mid h_1) = 0$ and $\Pr(p \cdot \sim q \mid h_2) = 0$. Then $\Pr(p \equiv q \mid h_1) = 1 - [\Pr(p \cdot \sim q \mid h_1) + \Pr(\sim p\ \cdot\ q \mid h_1)] = 1$ while similarly $\Pr(p \equiv q \mid h_2) = 1$.)

II

Let us begin, as one so often does in the philosophy of confirmation, with Hempel's classic 'paradox of confirmation' concerning empirical support for conditional generalities of form

(7) All As are Bs.

Since it is intuitively obvious—or so it seems—that (7) is confirmed by observing an A which has property B, i.e. by evidence of form $Aa \cdot Ba$, while by the same logic, evidence of form $\sim Ba \cdot \sim Aa$ confirms

$$\text{(8)} \qquad \text{All non-}B\text{s are non-}A\text{s},$$

why is it that observing some non-B to be a non-A feels confirmationally irrelevant to (7) when (7) and (8) are prima facie logically equivalent? I have previously explored this puzzle—which is a paradox of intuition only, not of logic—at some length (Rozeboom, 1968); but with the aid of principle (4a) its heart can be bared in a few sentences.

Consider first the alleged confirmation of (7) by 'positive instance' $Aa \cdot Ba$. Whatever may be the most appropriate technical reading of 'All As are Bs'—and as we shall see, this is not nearly so univocal as has often been assumed—it is analytically clear that (7) and the hypothesis that some particular object a is an A jointly entail that a is a B; hence given any background information k,

$$\text{(9)} \qquad \begin{aligned} \mathrm{CR}((7), Ba \mid Aa \cdot k) &= \frac{\Pr(Ba \mid Aa \cdot (7) \cdot k)}{\Pr(Ba \mid Aa \cdot k)} \\ &= \frac{1}{\Pr(Ba \mid Aa \cdot k)} \geq 1 \end{aligned}$$

in which the inequality is strict so long as Ba is not already certain given just Aa and k. Thus for natural k (specifically, which does not pre-empt the force of (7) for the observation in question), Ba is guaranteed to confirm (7) given k and Aa. This, however, is not quite the conclusion wanted here, for the new evidence with which observation of favorable object a augments our background information k is that a is both an A and a B. To justify our confirmational intuition in this case it needs be shown that 'All As are Bs' is confirmed under k by conjunctive datum $Aa \cdot Ba$, not just by datum Ba once Aa is also established.[2] Happily, principle (4a) makes completely clear the nature of this distinction. For by direct substitution we have

$$\text{(10)} \qquad \mathrm{CR}((7), Aa \cdot Ba \mid k) = \mathrm{CR}((7), Aa \mid k) \times \mathrm{CR}((7), Ba \mid Aa \cdot k).$$

Hence if datum Aa is by itself confirmationally irrelevant, given k, to whether all As are Bs—i.e. so long as $\mathrm{CR}((7), Aa \mid k = 1$—joint observation Aa and Ba confirms this generality to exactly the same degree, given k, as Ba confirms it given Aa as well as k. However, a greater-than-unity value for the second confirmation ratio on the right in (10) cannot prevent a sufficiently low value of the other from

[2] This point is well worth emphasis, for much of the literature on Hempel's paradox in fact reeks of failure to distinguish confirmation of (7) by Aa-and-Ba given k from its confirmation by Ba given Aa-and-k.

dragging their product below unity. Thus despite confirmation of (7) by Ba given Aa and k, the total epistemic import of $Aa \cdot Ba$ for (7) given k can indeed be disconfirmatory if Aa by itself disconfirms (7) given k.

But how, you ask, could observing just that something is an A possibly make any difference for whether all As are Bs—at least under natural background information which does not include some philosophical contrivance such as 'If anything is an A, then all As are Bs'? For we know from (2a) that Aa confirms (7) if and only if (7) confirms Aa, and surely the hypothesis that all As are Bs implies nothing about whether some object of yet-unidentified character will prove to be an A!? Your surprise nicely illustrates how profoundly we have allowed unexamined intuition to dominate our thinking on this matter. I do not suggest that this intuition is baseless and should be ignored. Quite the opposite: My contention here, as before (Rozeboom, 1968), is that it is an extrusion of our beliefs about nomic order which promises to yield important new analytical leverage on the latter. Even so, it is an intuition which turns out to be generally in error. This point, and its possible significance, is the main concern of this essay; but since I have already aroused the beast of Hempel's paradox, which others have so often sought to slay, I will pause long enough to pull its fangs before proceeding to the treasure which lies beyond.

Regardless of whether generalities (7) and (8) are entirely equivalent, the bite of Hempel's paradox lies in our intuition that finding a non-B which is not an A—a 'positive instance' of (8)—ought not to matter for whether all As are Bs. However, by principle (4a), the degree of (7)'s confirmation by datum $\sim Aa \cdot \sim Ba$ given background k factors is

$$\text{(11)} \qquad \mathrm{CR}((7), \sim Aa \cdot \sim Ba \mid k) = \mathrm{CR}((7), \sim Aa \mid k) \times \mathrm{CR}((7), \sim Ba \mid \sim Aa \cdot k).$$

If intuition is right to insist, as it does for natural k, that once something is known to be *not* an A its B-state has no further relevance to whether or not all As are Bs, then the last term in (11) is unity and observing that a is neither an A nor a B confirms (7) to the very same degree as does datum $\sim Aa$ alone—which, in turn, confirms or disconfirms (7) just in case Aa disconfirms or confirms it, respectively. Consequently, once it becomes clear that generalized conditionals are *not* as a rule confirmationally independent of data concerning just their antecedents, then it will no longer seem strange that $\sim Aa \cdot \sim Ba$ might confirm (7), or, alternatively, that $\sim Ba$ may be so disconfirmatory of all non-Bs being non-As that joint observation $\sim Ba \cdot \sim Aa$ does *not* confirm the latter and hence not that all As are Bs, either.

III

Why should I expect you to take seriously the patently silly proposal that whether or not all As are Bs is generally relevant to whether some particular thing is an A? Because once one stops to investigate the matter it soon becomes evident that this must be so. The deeper question is why intuition should be so insistent to the contrary and what special circumstances, if any, would make this intuition correct?

Consider first of all that insomuch as (7) and Aa jointly entail Ba for any object a, (7) by itself entails the extensional $Aa \supset Ba$. Now, having a false antecedent suffices—as is well known and often rued—to make an extensional conditional true. Since $\sim Aa$ thus confirms $Aa \supset Ba$, we are assured by principle (5) that Aa disconfirms it. But from principle (6a), since $\Pr(\sim [Aa \supset Ba] \mid (7) \cdot k) = 0$, (or from (4a), since (7) is equivalent to $(7) \cdot [Aa \supset Ba]$),

(12) $$\mathrm{CR}((7), Aa \mid k) = \mathrm{CR}(Aa \supset Ba, Aa \mid k) \times \mathrm{CR}((7), Aa \mid [Aa \supset Ba] \cdot k),$$

in which the first term on the right is, as just observed, less than 1, while the last term therein is unity if the relevance of (7) for Aa given k is mediated entirely by $Aa \supset Ba$. Hence, *unless* 'All As are Bs' *has some bearing on a's being an A over and above the import of $Aa \supset Ba$ for this, datum Aa disconfirms that all As are Bs via its disconfirmation of $Aa \supset Ba$.* Moreover, under the classic extensional reading of 'All As are Bs' as

(13) $$(x)(Ax \supset Bx),$$

it becomes extraordinarily difficult to see what relevance this generality *could* have, under natural k, for a's prospects on B except through its instantiation $Aa \supset Ba$ for a.

However, we need not insist that (13) captures the full ordinary-language force of 'All As are Bs'. Indeed, a good reason for *not* assuming this is simply that we do *not* normally consider data of form Aa to confirm this conditional generality. There are, in fact, at least two other prime candidates for the interpretation of all/are statements. One is to paraphrase (7) as

(14) If anything is an A, then it is a B,

which emphasizes universal quantification over a propositional connective and may hence be formalized as

(14a) $$(x)(Ax \to Bx),$$

and then to construe the if/then connective $\to$ as stronger than material implication. Alternatively, (7) may be understood as a statistical assertion of form

(15) The proportion of As which are Bs is r,

or

(15′) The probability of B-ness, given A-hood, is r,

both of which may be formalized as

$$\text{fr}(B \mid A) = r \tag{15a}$$

so long as the distinction between statistical probability and de facto relative frequency is not crucial for the purpose at hand. Schemata (14a) and (15a) differ in two important technical details. In the first place, (15a) contains a numerical parameter r whose unitary value in the statistical interpretation of (7) is merely one of alternative r-possibilities whose non-extreme cases we understand fully as well as we do $r = 1$; whereas to give (14a) comparable scope we would have to introduce a whole family of implicational connectives $\xrightarrow{r}$ such that for $r < 1$, Aa and $Aa \xrightarrow{r} Ba$ only make Ba likely to a certain degree. Secondly, generalizations of form (14a) entail stronger-than-extensional conditionalities between particular events, namely, that for every object a, the event of a's A-ing produces a B-ing by a.[3] In contrast, (15a) expresses a second-level relation between properties A and B which logically entails no conclusion about any particular object a except when $r = 1$, in which case (15a) implies only the extensional conditional $Aa \supset Ba$.[4] The differences among (13), (14a), and (15a) are sufficiently great that no conclusion about the confirmational behavior of one is certain to hold for the others as well. And if either (14a) or (15a) can be shown to be confirmationally independent of Aa, that one should clearly replace (13) as our top-drawer explication of (7) if we wish to preserve the latter's full intuitive force.

That (14a) may well have the confirmational property we seek is made plausible by recalling the sustained failure of philosophy of science to come up with a creditable extensional analysis for subjunctive and counter-factual conditionals. It is essential to the meanings of statements such as

> If the bullet had struck a quarter-inch lower, Jones would have been killed instantly,

and

> If Dull Victory wins the Derby this year, his stud fees will triple,

---

[3] Assertion by '$Aa \rightarrow Ba$' that an objective relation $\rightarrow$ holds between the events Aa and Ba is relatively straightforward so long as Aa is in fact the case, but how might '$Aa \rightarrow Ba$' then be counterfactually true (as we often think it is) when there is no such event as a's A-ing? Perhaps the ontologically safest way to interpret (14a) is as an ellipsis for $(x)[Ax \supset (Ax \rightarrow Bx)]$.

[4] Even this isn't strict deductive entailment unless *fr* is relative frequency (*contra* statistical probability) in a finite class, since otherwise $\text{fr}(B \mid A) = 1$ is technically compatible with the existence of a finite number of As which are not Bs.

that falsity of their antecedents not be a sufficient condition for their truth as a whole. But if refuting the antecedents of such conditionals is analytically forbidden to verify them, can the truth state of their antecedents make any confirmational difference for them at all? Certainly commonsense is adamant that knowing how Dull Victory's stud fees will be affected by the outcome of this year's Derby is totally useless for judicious betting on the race itself. When we assert 'If p then q in this sense, we mean that if p is/had been the case, then q is/would have been the case *because* of p. In particular, when this conditionality is thought to be causal,[5] 'If p then q' envisions p as bringing q about—i.e. *forcing* q to occur—by some principle of inter-event action.[6]

It might thus seem to be analytically true of the if/then connective in subjunctive and counterfactual conditionals that for any propositions p and q,

$$\Pr(q \mid p \cdot [p \rightarrow q] \cdot k) = 1 \tag{16}$$

for every k with which p and $p \rightarrow q$ are compatible, and

$$\Pr(p \mid [p \rightarrow q] \cdot k) = \Pr(p \mid k) \tag{17}$$

for most natural k, including in particular the null-background case where k is tautological. Indeed, this seemed reasonable enough to me previously (Rozeboom, 1968) to warrant defining a concept of 'modalic entailment' in terms of these properties. But unhappily, (17) proves to be far too tough a condition to hold in any significant generality—as I will now show not just for the conjectured propositional connective $\rightarrow$ but for any construction which has the confirmational properties ascribed to $\rightarrow$ by (16, 17).

Let H be a propositional schema, two different completions of which entail propositions H_{pq} and H_{qp}, respectively, which has if/then force in that for any two propositions r and s, r and H_{rs} jointly entail s. For example, if H is the schema '___$\rightarrow \ldots$', then H_{pq} and H_{qp} are '$p \rightarrow q$' and $q \rightarrow p$, respectively, while if H is the schema 'All ___s are ... s' or some proposed explication thereof, H_{pq} and H_{qp} might be 'If Aa then Ba' and 'If Ba then Aa' for object a.[7] (Heuristically, H_{pq} may be read as 'If p then q' or 'p brings q', but particular versions of H may endow H_{pq} with a much richer meaning than just this.) Our problem is now to

[5]As noted in Rozeboom (1968), the concept of 'because' subsumes logical dependencies as well as causal ones.

[6]Any reader who thinks that the notion of one event forcing another to occur is a primitive superstition long abandoned by modern science will find it edifying to browse through Chapter 1 of Blalock (1961)

[7]Note that propositions r and s are not required to be literally contained in H_{rs}. Thus when p and q are 'If Aa then Ba' and 'If Ba then Aa', respectively, H_{pq} and H_{qp} might respectively be just 'All As are Bs' and 'All Bs are As'. The important technical point here is that the present argument applies to the confirmational behavior of universal and particular conditionals alike.

assess how general may be the backgrounds k under which

(18) $$\Pr(p \mid H_{pq} \cdot k) = \Pr(p \mid k), \quad \Pr(q \mid H_{qp} \cdot k) = Pr(q \mid k),$$

or equivalently,

(18a) $$\mathrm{CR}(p, H_{pq} \mid k) = 1, \quad \mathrm{CR}(q, H_{qp} \mid k) = 1\ ,$$

obtain jointly. The point of the question of course is that if propositions p and q are essentially alike in their logical relations to k, then if k is a natural background under which p's truth state is intuitively irrelevant to whether or not p brings q, q should likewise be irrelevant under k to whether q brings p. This is especially so if k is null (i.e. tautological) while p and q have the same logical form.

Unfortunately, however, (18) holds only under constraints far too strong, even when k is null, to demand of a rational belief system in any generality. To begin, note that (18) directs us to consider jointly the two pairs of possibilities, that p may or may not bring q and that q may or may not bring p. Together, these yield four mutually exclusive and jointly exhaustive alternatives, $H_{pq} \cdot H_{qp}, H_{pq} \cdot \sim H_{qp}, \sim H_{pq} \cdot H_{qp}$, and $\sim H_{pq} \cdot \sim H_{qp}$, the first of which is apt to be analytically impossible if H embodies a strong sense of conditionally. (E.g., whereas $p \supset q$ and $q \supset p$ are clearly compatible for most p and q, p's being causally responsible for q presumably precludes q's being a cause of p.) Now: Does information about possibility H_{qp} upset the intuitive irrelevance of p to H_{pq}? Surely not. Yet if not, we are impaled on the dilemma which follows from Lemma 4 by substituting H_{pq} and H_{qp} for r and s, respectively— or worse, if H_{pq} and H_{qp} are incompatible, are spitted upon the single shaft which similarly protrudes from Lemma 4$'$:

Theorem 1: If $\mathrm{CR}(p, H_{pq} \mid k) = 1$ and $\mathrm{CR}(q, H_{qp} \mid k') = 1$ when k is variously null, H_{qp}, and $\sim H_{qp}$, and k' is variously null, H_{pq}, and $\sim H_{pq}$, then either $\mathrm{CR}(H_{pq}, H_{qp}) = 1$ or $\Pr(p \equiv q \mid H_{pq} \cdot \sim H_{qp}) = \Pr(p \equiv q \mid \sim H_{pq} \cdot H_{qp}) = 1$. (Nonzero: p, q, H_{pq}, H_{qp}, $\sim H_{pq}$, $\sim H_{qp}$, p/H_{pq}, q/H_{qp}, $\sim H_{qp}/H_{pq}, \sim H_{pq}/H_{qp}$).

Theorem 2: If H_{pq} and H_{qp} are incompatible while $\mathrm{CR}(p, H_{pq} \mid k) = 1$ and $\mathrm{CR}(q, H_{qp} \mid k') = 1$ when k is variously null and $\sim H_{qp}$, and k' is variously null and $\sim H_{pq}$, then $\Pr(p \equiv q \mid H_{pq} \cdot \sim H_{qp}) = \Pr(p \equiv q \mid \sim H_{pq} \cdot H_{qp}) = 1$. (Nonzero: p, q, H_{pq}, H_{qp}, $\sim H_{pq}$, $\sim H_{qp}$, p/H_{pq}, q/H_{qp}).

Theorems 1 and 2 tell us that intuitive principle $\mathrm{CR}(p, H_{pq} \mid k) = 1$ can be had for the choices of k listed therein only at a price—or, if H_{pq} and H_{qp} are not mutually exclusive, at one of two prices. The price $\Pr(p \equiv q \mid H_{pq} \cdot \sim H_{qp} = \Pr(p \equiv q \mid \sim H_{pq} \cdot H_{qp}) = 1$ which is our only option for strong senses of conditionality, is totally unacceptable, for it says that given if-p-then-q but not if-q-then-p or conversely, then p if and only if q. This would be self-contradictory were the if/then force of H merely extensional implication, and is in any event an absurd requirement to impose on a belief system. But even for weak if/thens which allow the alternative option in Theorem 1, $\mathrm{CR}(H_{pq}, H_{qp}) = 1$ does not seem like a reasonable stricture either; for even if H_{pq} and H_{qp} are compatible, why should information about the one *inherently* make no difference for the credibility of the other? It may be concluded, then, that intuitive principle $\mathrm{CR}(p, H_{pq} \mid k) = 1$ can be defended even for null k only by abandoning it when k is H_{qp} or $\sim H_{qp}$.

But is this a reasonable retrenchment? For while p should be more likely given H_{qp} than given $\sim H_{qp}$, it is extremely difficult to see how H_{pq} could have any more confirmational bearing on p once H_{qp} or $\sim H_{qp}$ is given than it does in the absence of the latter information. This is especially true of $\sim H_{qp}$, which by Theorem 2 is all that matters for strong if/thens. If knowledge that p brings q tells nothing about whether p is the case, how does information that q does not bring p disturb that irrelevance? But if intuition can be this wrong about $\mathrm{CR}(p, H_{pq} \mid k)$ when k is $\sim H_{qp}$, dare we trust it at all on this point?

Moreover, we cannot restore (18) to confirmation-theoretic health just by relinquishing the intuitive indifference of p to H_{pq} given H_{qp} or $\sim H_{qp}$, for even then (18) remains incompatible with many credibility combinations which we have no present reason to consider irrational. To show this in an extreme case, suppose that $\Pr(H_{pq} \cdot H_{qp}) = \Pr(\sim H_{pq} \cdot \sim H_{qp}) = 0$, i.e. that H_{pq} and H_{qp} are thought to be jointly exhaustive as well as mutually exclusive. Then $\Pr(p \mid H_{pq}) = \Pr(p)$ and $\Pr(q \mid H_{qp}) = \Pr(q)$ jointly imply $\Pr(p \mid H_{pq}) = \Pr(p \mid H_{qp})$ and $\Pr(q \mid H_{pq}) = \Pr(q \mid H_{qp})$, from which in turn, since $\Pr(p \cdot \sim q \mid H_{pq}) = \Pr(\sim p \cdot q \mid H_{qp}) = 0$, it follows (cf. proof of Lemma 4) that $Pr(p \equiv q \mid H_{pq} \cdot \sim H_{qp}) = \Pr(p \equiv q \mid \sim H_{pq} \cdot H_{qp}) = \Pr(p \equiv q) = 1$. Yet why should one who is convinced that either H_{pq}-or-H_{qp} is the case be required to deny all possibility of $p \cdot \sim q$ or $q \cdot \sim p$ when neither of these two precluded alternatives is at all incompatible with $H_{pq} \cdot \sim H_{qp} \vee \sim H_{pq} \cdot H_{qp}$?

One can idle away many a cheerful hour in search of unusually outrageous constraints which follow from (18), especially if still another intuition almost as strong as (18), namely, that H_{pq} should not affect the credibility of q given not-p, is also thrown into the game; and I will not totally foreswear a future return to this matter. Meanwhile, until we learn much more about what, over and above axiomatic coherence (i.e. agreement with the axioms of the probability calculus) determines the rationality of a belief system, hoping to find some sense of conditionality with built-in confirmational indifference to its antecedent is like yearning for the gold at rainbow's end. A proposition which implies that if-p-then-q *may* be confirmationally irrelevant to p in particular instances, but only because certain other credibilities in the system, notably those involving the parallel possibility that if-q-then-p, oddly happen to hit upon just the right numerical values. In general, then, we must assume that $\Pr(p \mid k$ and if-p-then-$q)$ is *not* equal to $\Pr(p \mid k)$, not even for null k, nor can this lack of indifference be expunged by a sufficiently ingenious contrivance of new conditionality concepts.

But what, then, are we to make of the intuition which insists so strongly that q's being conditional upon p has no relevance to whether p is the case? Is it just a witless blunder which deserves no further heed? I think not. It is at the very least a philosophically profound (*contra* silly) error which, sedated against naive overexuberance, may yet be usefully rehabilitated.

There are perhaps two reasons, one a superficial reflection of the other, why it is so tempting to consider if-p-then-q indifferent to p. Superficially, this is a failure to distinguish logical independence from informational irrelevance. Unlike the relation between an extensional conditional and its antecedent, whose negations are incompatible, the possibility envisioned by a subjunctive if-p-then-q has no logical relevance to whether p is the case. Hence it is no easier for common sense to appreciate that if-p-then-q might inform about p than it can see how learning that John has red hair could affect the credibility that John loves Mary—one's natural inclination in both cases being to assume that there is no confirmational relevance at all. However, logical and confirmational independence are *not* the same, neither analytically nor extensionally. It is quite appropriate for our learning of John's redheadedness to modify the credence we give to his loving Mary if e.g. we suspect that redheads are more passionate than other people; and knowing that p brings q would indeed be evidence for p were it believed that whatever establishes subjunctive dependencies also tends to actualize their antecedents.

But I am being unfair to intuition here. My examples point out that there may well be natural background information under which logically independent propositions are confirmationally related. But common sense does not dispute that; rather, the confirmational-independence intuition becomes ascendent only with an increasing impoverishment of background information under which the logical possibilities for one state of affairs become increasingly symmetric in their relations to the possibilities for another. In the limiting case of *no* background information, then, perhaps logical and confirmational independence do coincide after all, at least extensionally. I venture that the deep reason why (18) seems so plausible is simply that for null k it is an instance of

The Fundamental Indifference Intuition [FII]. If propositions p and q are logically independent, then $\mathrm{CR}(p, q \mid k) = 1$ if k is null.

However, even apart from qualms about whether credibility relations are really well-defined at all relative to no background whatsoever, it is important to realize that *FII* is dramatically untenable so long as 'logical independence' is understood in its usual sense whereby two or more propositions are logically independent (of one another) if all their truth-combinations are logically possible. Let p and q be any two logically independent propositions while $r =_{\text{def}} p \cdot q \vee \sim p \cdot \sim q$. Then any two of the three propositions p, q, r are logically independent of one another (since for p and r, $p \cdot r$, $p \cdot \sim r$, $\sim p \cdot r$, and $\sim p \cdot r$ are logically equivalent to $p \cdot q$, $p \cdot \sim q$, $\sim p \cdot q$, and $\sim p \cdot q$, respectively, and similarly for r's logical independence of q), so by *FII*,

$$\text{(19)} \qquad \Pr(p \cdot q) = \Pr(p) \times \Pr(q), \quad \Pr(p \cdot r) = \Pr(p) \times \Pr(r) \\ \Pr(q \cdot r) = \Pr(q) \times \Pr(r).$$

From (19), letting $w = \Pr(p \cdot q)$, $x = \Pr(p \cdot \sim q)$, $y = \Pr(\sim p \cdot q)$, $z = \Pr(\sim p \cdot \sim q)$, we have

$$\text{(20)} \qquad w = (w + x)(w + y),\ w = (w + x)(w + z), \\ w = (w + y)(w + z),$$

which a little algebra shows to obtain if and only if either w, x, y, or z is unity or $w = x = y = z$, i.e. only if either (*i*) $\Pr(p)$ and $\Pr(q)$ are both extreme (i.e. zero or unity) or (*ii*) $\Pr(p) = \Pr(q) = 0.5$. But a credibility system which satisfies either (*i*) or (*ii*) for every pair of logically independent propositions is stunningly degenerate. (The degeneracy of (*i*) is obvious; while for a system containing three logically independent propositions p, q, and r, since r is then also logically independent of $p \cdot q$ and $p \cdot \sim q$, (*ii*)'s holding for all of these would require $\Pr(p) = \Pr(p \cdot q) = \Pr(p \cdot \sim q) = 0.5$, which is impossible.) Hence in general, $\text{CR}(p, q) \neq 1$ even when p and q are logically independent, and not merely does our intuitive skill at recognizing confirmational relevance fail in relatively esoteric cases like (18), it is demonstrably rotten at the root.

Or is it? May there not exist some reading of 'logical independence', stronger than its technically standard sense, under which *FII* is tenable after all? (For example, the conceptual linkage between p and $p \cdot q \vee \sim p \cdot \sim q$ in my disproof of *FII* might seem to violate the spirit if not the letter of 'logical independence', though to be sure if *FII*'s interpretation of this were to allow no conceptual overlap at all between logically independent propositions it would no longer explain the intuitive confirmational independence between a subjunctive conditional and its antecedent.) Indeed, unless it is arbitrary which propositions in a belief system are confirmationally independent under null background information, this must be so. For if there is some *basis* for propositions p and q being confirmationally independent under null background while p and r are not, even though there are no more logical exclusions on truth-combinations of p and q than of p and r, then a strong sense of 'logical independence' which yields *FII* can be defined on this basis. Contrapunctally, if a person considers p and q, but not p and r, to be confirmationally independent under null background, then, unless such opinions are rationally a matter of whim, that person has in effect commited himself to some possibility for how things are, namely, to whatever would make his particular allocation of confirmational independencies an appropriate one.

In short, the situation is this: Given a boolian algebra $\{p_i\}$ of propositions, there are infinitely many different unconditional credibility distributions $\{\mathrm{Pr}(p_i)\}$ over these which are 'formally admissible' in that they satisfy the restrictions of the probability calculus. Are all these admissible credibility allocations equally rational, or do some have greater epistemic merit than others? Nothing in the confirmation-theoretic literature, whether by philosophers, statisticians, or practicing scientists, has yet established any convincing grounds on which to prefer one admissible distribution of unconditional credibilities to another.[8] Yet were this completely arbitrary, then so would be the conclusion of any ampliative argument; for if premise p neither entails nor is incompatible with prospective conclusion q given background information k, then for any nonnegative real number n there exists an admissible unconditional credibility distribution under which $\mathrm{CR}(p, q \mid k) = n$. Accordingly, if the unconditional credibilities in a person's belief system were entirely free parameters, normatively constrained only by formal admissibility, then apart from logical entailments the degree to which any given proposition confirms another would likewise be arbitrary. Yet surely, when I think that p confirms q given k while you think that disconfirms it, we have a genuine cognitive dispute rather than just a mismatch of personal tastes in free parameters. Even if your credibility system agrees with mine in rating $\mathrm{Pr}(p \cdot q \cdot k) \times \mathrm{Pr}(k)$ higher than $\mathrm{Pr}(p \cdot k) \times \mathrm{Pr}(q \cdot k)$, so that for both of us $\mathrm{CR}(p, q \mid k) > 1$, I may still wonder if we have *correctly* construed p to confirm q given k. Is it then a mistake for me to feel normative qualms about whether the structure of uncertainty in my belief system—i.e. what evidence confirms what conclusions under what background information—is exactly as it *should* be even though my credences are

[8] There are at least two important partial exceptions to this which may be expected to play an increasingly major role in confirmation theory but are still afflicted with internal difficulties. One is that symmetry properties of one sort or another have often seemed to be a reasonable demand on unconditional credibilities—i.e. that if p and q are formally alike in certain ways, then $\mathrm{Pr}(p)$ ought to equal $\mathrm{Pr}(q)$. Unfortunately, symmetry stipulations tend to entail inconsistencies unless carefully restricted (cf. the vicissitudes of the classic 'principle of insufficient reason'), and just what symmetry demands can be successfully defended under what boundary conditions is still very much an open question. Secondly, applications of statistical generalities to inferences about particular events usually presuppose that some event credibilities conditional upon statistical information are analytically determinate, thereby placing constraints beyond formal admissibility on which unconditional credibilities are rationally acceptable. For example, it is usually assumed that the credibility of a particular object a's being an A, given that the relative frequency or statistical probability of A-hood is r, is r; whence it follows that $\mathrm{Pr}[Aa \cdot \mathrm{fr}(A) = r] = r \times \mathrm{Pr}[\mathrm{fr}(A) = r]$. But this principle, too, is considerably more problematic than it intuitively appears (cf. Rozeboom, 1969).

in perfect accord with the constraints of the probability calculus? Is there not *something* which, could I but become cognizant of it, would reassure me/confirm my suspicions that my uncertainty structure is satisfactory/misguided?

I do not pose these questions rhetorically, for the proper directions in which to answer them seem far from clear to me. I do, however, insist that the questions themselves are extraordinarily important ones whose full depths and complexities will unfold only gradually as we tease apart the issues which coil within them. Previously, I have ventured that placements of confirmational independence within a person's belief system are intimately tied to his suppositions about the world's responsibility order (1968), and the final consideration to be explored here will provide powerful reinforcement for that conjecture along lines rather different from my earlier argument.

IV

We have yet to probe the confirmational import of datum Aa for generalization 'All As are Bs' when this is interpreted statistically. Section III has already established (contrary to my intimations in (1968)) that the credibility of Aa cannot be generically indifferent to information about the frequency of B-ness among things which are As, but it will nonetheless prove instructive to examine the specifics of this case.

To expedite the discussion, I had best review a few technical concepts which will be needed. A scientific *variable* $\mathbf{X}$ over a domain or 'population' P of objects is a function from P into a set $\{X_i\}$ of attributes (or, if $\mathbf{X}$ is numerically scaled, into numbers which represent attributes) such that for each object a in P exactly one attribute in $\{X_i\}$ applies to a, this attribute being what is meant by the *value* of variable $\mathbf{X}$ for argument a. (For details, see (1961, 1966).) For conceptual simplicity—and technically, this is an *enormous* simplification which, happily, is entirely harmless for present purposes—I shall speak as though a variable has only a finite number of alternative values.

The statistical *distribution* of variable $\mathbf{X}$ in population P is a function $\mathrm{fr}(\mathbf{X})$ which maps each value X_i of $\mathbf{X}$ into the relative frequency or statistical probability $\mathrm{fr}(X_i)$ of attribute X_i in P, while more generally the joint distribution of two (or, similarly, more) variables $\mathbf{X}$ and $\mathbf{Y}$ in P is a function $\mathrm{fr}(\mathbf{XY})$ mapping each combination of a value X_i of $\mathbf{X}$ with a value Y_j of $\mathbf{Y}$ into the relative frequency or statistical probability $\mathrm{fr}(X_iY_j)$ in P of the conjunction of attributes X_i and Y_j. The *conditional* distribution of $\mathbf{Y}$ given $\mathbf{X}$ in P is a function $\mathrm{fr}(\mathbf{Y} \mid \mathbf{X})$ mapping each pair of values X_i of $\mathbf{X}$ and Y_j of $\mathbf{Y}$ into the relative frequency or statistical probability $\mathrm{fr}(Y_j \mid X_i)[= \mathrm{fr}(X_iY_j)/\mathrm{fr}(X_i)]$ of Y_j among just those members of P which have attribute X_i, while the conditional distribution of $\mathbf{Y}$ in P given a particular value X_i of $\mathbf{X}$ is $\mathrm{fr}(\mathbf{Y} \mid \mathbf{X})$ restricted to X_i.

Very often what is of statistical interest about a given distribution is not so much its unanalyzed entirety as certain of its parametric features such as means, variances, correlations, etc. If D is the family of alternative distributions logically possible for a given set of variables in a certain population, then a *parameter* of D is any many-one function from D into numbers (or sometimes more complex abstract entities), while a *parameterization* of distribution family D is a set of parameters such that each distribution in D is uniquely identified by the totality of its values on these parameters. (Henceforth, when I speak of a distribution 'family' I shall always mean the set of alternative distribution possibilities for a fixed set of variables in a particular population.) There are many alternative ways to parameterize a given distribution family; however, if $\langle \xi, \Omega, \ldots \rangle$ and $\langle \xi', \Omega', \ldots \rangle$ are two different parameterizations of D then each parameter in the one set is completely specified by the parameters in the other, i.e. for the first parameter in the second set (and similarly for the others) there exists, a function ϕ such that $\xi'(d) = \phi[\xi(d), \Omega(d), \ldots]$ for each d in D. Parameters $\xi,\ \Omega, \ldots$ of distribution family D are all *logically independent* (of one another) iff any logically possible value on any one of these parameters is logically compatible with all logically possible combinations of values on the others.

Finally, a notational ellipsis: When discussing our uncertainty about the numerical values of a given distribution's parameters, it will be convenient to let 'The value of parameter ξ for distribution d is ξ_i', be abbreviated simply as 'ξ_i'. Then $\Pr(\xi_j \mid \Omega_i)$ is the credibility that the joint distribution at issue has value ξ_j of parameter ξ given that its value of parameter Ω is Ω_i.

When 'All As are Bs' is interpreted statistically as the claim that $\mathrm{fr}(B \mid A) = 1$, the theory of its confirmation by instances falls under the more general problem of determining the value of a statistical parameter from sample data. In our present case, the distribution at issue is that of two dichotomous variables **A** and **B**, whose values are A, $\sim A$ and B, $\sim B$, respectively, in some population which may here be described noncommitally as the class T of 'things'. Because these variables are both dichotomies, their joint distribution in T can be exhaustively specified by just three numbers, two such parameterizations being

$$(21) \qquad \alpha =_{\mathrm{def}} \mathrm{fr}(A), \quad \beta =_{\mathrm{def}} \mathrm{fr}(B \mid A), \quad \gamma =_{\mathrm{def}} \mathrm{fr}(B \mid \sim A),$$

and

$$(22) \qquad \alpha' =_{\mathrm{def}} \mathrm{fr}(B), \quad \beta' =_{\mathrm{def}} \mathrm{fr}(A \mid B), \quad \gamma' =_{\mathrm{def}} \mathrm{fr}(A \mid \sim B).$$

(In what follows, 'α', 'β', etc. will be used specifically as defined in (21) and (22).) Each parameter in (22) is a determinate function of the parameters in (21), namely,

$$\alpha' = \alpha\beta + (1-\alpha)\gamma, \qquad \beta' = \frac{\alpha\beta}{\alpha\beta + (1-\alpha)\gamma}\,,$$
$$\gamma' = \frac{\alpha(1-\beta)}{(1-\beta) + (1-\alpha)(1-\gamma)}$$

and similarly with sets (21) and (22) interchanged. *Within* each set (21) or (22), however, the parameters are all logically independent of one another. (Still another parameterization of this distribution is

$$\mathrm{fr}(A \cdot B), \qquad \mathrm{fr}(A \cdot \sim B), \qquad \mathrm{fr}(\sim A \cdot B).$$

These last are not logically independent, however, insomuch as 1 less the value of any one is an upper bound on the value of any other.)

What does observing that something is an A tell us about the joint distribution of $\mathbf{A}$ and $\mathbf{B}$ in T? Or rather, to emphasize our present special interest, when is a parameter ξ of $\mathrm{fr}(\mathbf{AB})$ *not* informed about by datum Aa—i.e. under what circumstances is it the case that for every value ξ_i of ξ, $\mathrm{CR}(\xi_i, Aa \mid k) = 1$? (Henceforth, unless there is need for explicit mention of background information, I will let k be null.) One parameter to which Aa clearly ought to be confirmationally relevant is α, i.e. $\mathrm{fr}(A)$, and indeed, under the orthodox assumption that $\Pr(Aa \mid \alpha_i) = \alpha_i$, it can easily be shown that

$$\mathrm{CR}(Aa \mid \alpha_i) = \frac{\alpha_i}{\mathrm{Exp}(\alpha)}\,,$$

where $\mathrm{Exp}(\alpha)$ is the mean of the credibility distribution over possible values of α. More useful for present purposes, however, is an assumption which will let us get at Aa's significance for the rest of $\mathrm{fr}(\mathbf{AB})$, namely, that the relevance of datum Aa for any parameter ξ of $\mathrm{fr}(\mathbf{AB})$ is mediated by whatever Aa tells about the frequency of As. Specifically, I shall assume that for all values α_i of α and ξ_i of ξ, Aa adds nothing to what α_i tells about ξ_j, i.e. that $\Pr(\xi_j \mid \alpha_i \cdot Aa) = \Pr(\xi_j \mid \alpha_i)$ or, equivalently,

(23) $$\mathrm{CR}(\xi_j, Aa \mid \alpha_i) = 1 \qquad \text{(assumed)}$$

(Adoption of principle (23)—which would have to be hedged against certain quirky special cases were it to be defended in complete generality is strictly provisional here, but as a working idealization it seems reasonable in ways which can be verbalized, and it is hard to think of an alternative which makes any intuitive sense.) From (23) by (6a) we then have

(24) $$\mathrm{CR}(\xi_j, Aa) = \sum^{i} \Pr(\alpha_i \mid Aa) \times \mathrm{CR}(\xi_j, \alpha_i)\,,$$

in which summation is over all possible values of α. Equation (24) says that the degree to which datum Aa confirms that fr(**AB**) has a particular value ξ_j of parameter ξ is a weighted average of the degrees to which the various possible values of α respectively confirm ξ_j. If parameters α and ξ are confirmationally independent of one another—i.e. if $\mathrm{CR}(\xi_j, \alpha_i) = 1$ for all α_i and ξ_j—then $\mathrm{CR}(\xi_i, Aa) = 1$. If α and ξ are *not* confirmationally independent, it is still possible for the right-hand side of (24) to equal unity, but only under special allocations of credibility which, so far as we have any reason to suspect, obtain in only a vanishingly small proportion of rational credibility systems.[9] Hence,

THEOREM 3. A parameter ξ of the joint distribution of variables **A** and **B** in population T is confirmationally independent of datum Aa if and, for all practical purposes, only if ξ and $\mathrm{fr}(A)$ are confirmationally independent of each other.

In particular, letting arbitrary parameter ξ be β, we have

COROLLARY. $\mathrm{fr}(B \mid A)$ is confirmationally independent of Aa if and, for all practical purposes, only if it is confirmationally independent of $\mathrm{fr}(A)$.

Under assumption (23), whether observing an A has any relevance to whether statistically all As are Bs thus reduces to whether information about the frequency of A-hood makes any confirmational difference for the conditional frequency of B-ness among things which are As. This, in turn, falls under the more general question of which statistical parameters are confirmationally independent of which others. So by rights, to achieve resolution on this point it should suffice to consult what advanced statisticians have had to say about it. However, extant technical doctrine on credibility relations among statistical parameters can be summarized with shocking brevity: *There is none.* Or rather, the theory of this exists only implicitly in real-life statistical practices which are as prevalent as they are unreasoned, namely,

The Statistical Independence Presupposition [SIP]. If parameters $\xi, \mathit{\Omega}, \ldots$ of a given distribution family are logically independent under background information k, then $\xi, \mathit{\Omega}, \ldots$ are also confirmationally independent given k.

It is not profitable to attempt documenting the universality of this presupposition

[9] Namely, when there is zero covariance between the two series of quantities $\{\mathrm{CR}(\xi_j, \alpha_i)\}$ and $\{\Pr(\alpha_i \mid Aa) - \Pr(\alpha_i)\}$ for $i = 1, \ldots$ This is mathematically possible, but has virtually no chance of being exactly true unless one of these two series has zero variance—while moreover the variance of the second series is zero only if parameter α is already completely known, i.e. if $\Pr(\alpha_i)$ equals 1 for one value of α and 0 for the rest.

here,[10] but readers familiar with the practice of partitioning sample data as a conjunction of 'sufficient' statistics, each of whose likelihood function depends only upon a proper subset of the unknown parameters, will also recall that sample values of these sufficient statistics are construed to inform *only* about their own specific likelihood parameters. As a less sophisticated illustration, a researcher who searches his journals for some indication of the correlation between certain variables would feel totally unenlightened if he could find reports only of their means and variances.

Yet no matter how attractively *SIP* may intuit in particular cases, it is altogether untenable as a general principle. For it can easily be proved that if $\xi, \Omega \ldots$ and $\xi', \Omega', \ldots$ are nontrivial[11] alternative logically independent parameterizations of the same distribution family, and parameters $\xi, \Omega \ldots$ are all confirmationally independent of one another, then this is *not* generally also true of $\xi', \Omega' \ldots$. In particular, this holds for alternative parameterizations (21) and (22) of fr(**AB**). To show this by means of an extreme example, suppose that our belief system gives nonzero and, say, equal credibility to only two values, 0.2 and 0.8, of α, and that the same is also true of β and γ. Then there are eight nonzero-credibility alternatives for fr(**AB**), one for each row in Table I, all having the same probability

TABLE I

Values of parameters (21) and (22) under eight different possibilities for fr(**AB**)

α	β	γ	α'	β'	γ'
0.2	0.2	0.2	0.20	0.20	0.20
0.2	0.2	0.8	0.68	0.06	0.50
0.2	0.8	0.2	0.32	0.50	0.06
0.2	0.8	0.8	0.80	0.25	0.25
0.8	0.2	0.2	0.20	0.75	0.75
0.8	0.2	0.8	0.32	0.50	0.94
0.8	0.8	0.2	0.68	0.94	0.50
0.8	0.8	0.8	0.80	0.80	0.80

[10] The only significant exceptions to it of which I know are Raiffa and Schaiffer (1961), who at least recognize the possibility of confirmationally related parameters in their abstract development even if they say nothing about it subsequently, and two important applications of correlated-parameter notions by Whittle (1957, 1958) to which Professor L. J. Savage has graciously called my attention. I think it can safely be said (i) that only Bayesian statisticians have shown awareness that *SIP* is not a statistical truism, while (ii) even Bayesians have remained at a loss how to replace it with some more defensible normative principle.

[11] I.e., the parameters in the one set do not have a one-one equivalence to those in the other, as would be the case e.g. if $\xi' = \xi, \Omega' = \Omega$, etc.

of 0.125 if α, β and γ are confirmationally independent. If the confirmation ratios $\mathrm{CR}(\alpha'_i, \beta'_j)$, $\mathrm{CR}(\alpha'_i, \gamma'_k)$, and $\mathrm{CR}(\beta'_j, \gamma'_k)$ are worked out for all values of α', β', and γ' which appear in the last three columns of Table I, these are seen to be either zero (complete disconfirmation), 4, or 8 (both strong and in some cases complete confirmation.) When the marginal credibility distributions for α, β and γ are more realistically continuous the confirmational relations among α', β', and γ' under confirmational independence of α, β and γ are too complex (at least for me) to evaluate analytically. However, computer-assisted finitary approximations indicate that even when the marginal credibilities over α, β and γ are perfectly flat (i.e. all their possible values are equally likely), the average deviation of the $\mathrm{CR}(\alpha'_i, \beta'_j)$, $\mathrm{CR}(\alpha'_i, \gamma'_k)$, and $\mathrm{CR}(\beta'_j, \gamma'_k)$ from unity is about 0.2, while if two of the three parameters in (21) are known fairly precisely when the third is rather uncertain, the confirmational relations among α', β', and γ' become quite strong.[12] It is, in fact, a nice question whether any nondegenerate distribution of credibilities over the possibilities for $\mathrm{fr}(\mathbf{AB})$ permits confirmational independence among parameters (21) to be combined with the same for parameters (22).

"Very well", I can hear you mutter, "so logically independent statistical parameters are not in general also confirmationally independent. What of it?" Don't turn off too quickly, for there is indeed much to be made of this. First and least important, it substantiates my earlier claim that the $\mathrm{fr}(B \mid A) = 1$ reading of 'All *A*s are *B*s' will not yield systematic confirmational indifference of conditional generalities to instances just of their antecedents. Secondly, it proposes a contribution to real-life statistical technology: Since in practice we always act as though estimates of one statistical parameter tell us nothing about the others which interest us when these are logically independent of the first under our background assumptions, we have not been reaping the full statistical harvest provided by our sample data and should revise our statistical models to exploit the parameter correlations which, we now see, must lurk somewhere within our belief system if it is formally admissible. Conceivably, we could even parlay these improved statistical methods into more powerful or more economical research designs. (It would, for example, save an *awful* lot of experimental effort if the regression of one variable upon another could be reliably estimated from the distribution of the latter alone.)

[12] Average CR deviation from unity, though easy to comprehend, is not actually a very good measure of confirmational relatedness. Much superior to it, technically, are the Information-theoretic measure of transmitted information and the Correlation Ratio for goodness of curvilinear regression. In terms of the latter measures, my finitary approximation shows the predictability of β' or γ' from α' and conversely to be negligible when the credibilities over α, β, and γ are flat (a correlation ratio of only 0.16, or 2% variance reduction), and to remain generally low even when α, β, and/or γ become known more precisely. In contrast, the correlation ratios for predicting β' from γ' and conversely are about 0.36 (a variance reduction of 13%) even when α, β, and γ have flat marginal credibilities, and become quite respectable in size when one or more of these marginal distributions is sharp.

However, I must in all honesty confess that any hopes I might promulgate for so enhancing our statistical engineering prowess would be essentially bogus. There are reasons to doubt that confirmational dependencies among logically independent statistical parameters are often large enough for practical significance except under conditions to which applied statisticians already accommodate intuitively through the background assumptions of the models they adopt on particular occasions.[13] Even so, it should be of considerable value for advanced statistics to work out the likely magnitude of these dependencies under various idealized and realistic circumstances, and more generally to think through the *logic* of confirmationally correlated statistical parameters for whatever new wisdom this may bring to the theory and practice of inductive inference.

An essential part of the aforementioned logic will be some normative standards concerning which parameters of what distributions *should* be confirmationally independent under a given informational background. For if I take the first of two different parameterizations of the same distribution family to be confirmationally independent under background k while for you this seems true of the second set, how is our disagreement to be appraised? Is it just a misalignment of arbitrary whims, or are we not each making some cognitive commitment in which at least one of us is wrong? Obviously much needs to be said on this matter, a great deal of it requiring considerable sophistication in statistics and experimental design, and needing first of all to grope out some sense of direction in those many sectors of the problem where even professionally sensitized intuition still flounders perplexedly. Even so, regarding what is probably the most basic of all ways to partition a joint statistical distribution, educated intuition does seem to me to speak with some assurance even if it takes a moderately trained ear to discern its message. This is

The Causal Structure Criterion [CSC]. With some qualifications, every parameter of $\mathrm{fr}(\mathbf{X})$ in population P is confirmationally independent of every parameter of $\mathrm{fr}(\mathbf{Y} \mid \mathbf{X})$ in P under background information k if and, for all practical purposes, only if k entails that variable $\mathbf{X}$ has complete causal precedence over variable $\mathbf{Y}$ in P. By '$\mathbf{X}$ has complete causal precedence over $\mathbf{Y}$ in P' is meant that any causal antecedents common to $\mathbf{X}$ and $\mathbf{Y}$ affect the variance of $\mathbf{Y}$ in P only through the mediation of $\mathbf{X}$, i.e. that any causal coupling between $\mathbf{X}$ and $\mathbf{Y}$ not held constant in P is directed from $\mathbf{X}$ to $\mathbf{Y}$ rather than from $\mathbf{Y}$ to $\mathbf{X}$ or to $\mathbf{X}$ and $\mathbf{Y}$ jointly from a mutual source.

In particular, for our two dichotomous variables $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A}$ has complete causal

[13]Notably, it is virtually always assumed that the to-be-estimated distribution is of a particular restricted form, given which a small subset of what would otherwise be logically independent parameters precisely determines the remainder. Thus if a multivariate distribution is assumed to be Normal, all its parameters can be computed from just its first and second moments, while the totality of a Poisson distribution is specified by its mean alone.

precedence over $\mathbf{B}$ iff our uncertainty about the frequency of $\mathbf{A}$ is independent of our uncertainty about the frequency of $\mathbf{B}$ both among things that are As and things that are non-As.

I am not going to discuss *CSC* in any detail here, much less try to justify it, for I voice it not as any sort of *conclusion* but merely as a scientifically realistic point of departure for what needs to be an extended series of probes and exchanges on possible relations between the confirmational grain of our statistical beliefs and our suppositions about nomic order. I describe *CSC* as 'scientifically realistic' because it embodies two powerful undercurrents in professional thinking on multivariate analysis and research design, namely, (i) that some parameterizations of a statistical distribution align more closely with its causal structure than do others, and (ii) that features of the world sharing no common determinants whatsoever are independent in some fashion stronger than mere logical independence. From (i), we have that $\mathrm{fr}(\mathbf{Y} \mid \mathbf{X})$ is a pure reflection of $\mathbf{Y}$'s causal dependence upon $\mathbf{X}$ (in both its character and degree) only if $\mathbf{Y}$ is not, conversely, a source of $\mathbf{X}$ and no additional sources of $\mathbf{Y}$ unmediated by $\mathbf{X}$ also influence $\mathbf{X}$. Otherwise, the statistical conditionality of $\mathbf{Y}$ upon $\mathbf{X}$ conflates $\mathbf{X}$'s causal import for $\mathbf{Y}$, if any, with $\mathbf{Y}$'s noncausal predictive regression on $\mathbf{X}$ in ways that, for linear dependencies at least, can be formulated in considerable quantitative detail. Secondly, if the sources of $\mathbf{X}$ and the sources of $\mathbf{Y}$'s statistical dependence on $\mathbf{X}$ have nothing in common—as presumably obtains if $\mathrm{fr}(\mathbf{Y} \mid \mathbf{X})$ reflects purely $\mathbf{X}$'s causal influence on $\mathbf{Y}$—then undercurrent (ii) presses us to construe $\mathrm{fr}(\mathbf{X})$ and $\mathrm{fr}(\mathbf{Y} \mid \mathbf{X})$ as strongly independent in some sense, the most promising candidate for this being confirmational independence.[14]

In short, *CSC* is what results if we attempt to salvage statistical intuition *SIP* by limiting it to just those statistical parameters which align without interpretive contamination with the causal structure of the distribution they describe. Just how robustly *CSC* can be maintained I have no present idea. Certainly it needs some restrictions on the background information over which it generalizes, insomuch as for any background k under which two statistical parameters ξ and $\mathit{\Omega}$ are confirmationally independent, additional information h can always be conceived such that ξ and $\mathit{\Omega}$ are no longer confirmationally independent given both k and h. Were this the only problem here, it might suffice to restrict *CSC* to just those k under which the parameters at issue are logically independent, since that will block the construction I have in mind while preserving the spirit of *SIP*; however, I can also think of other ways in which *CSC* may be overextended as given. But

[14] (ii)'s most common embodiment is in the classic and still widespread assumption that source variables which themselves have no deeper sources in common are statistically independent of one another. Statistical independence is not applicable to the present case, however, insomuch as the dependence of $\mathbf{Y}$ upon $\mathbf{X}$ in P is a parametrically constant *feature of* the joint distribution in P of $\mathbf{X}$, $\mathbf{Y}$, and other variables, not a variable *within* this system.

CSC's boundary limitations are not all that crucial just now. More important is for us to begin tuning our thoughts to the qualitative possibilities of principles such as this, to speculate whether *something* like it may not be an inevitably basic feature of any complex rational belief system.

SUMMARY AND PROSPECTUS

There is no good place to stop in a domain of new ideas so fragmentarily explored as this one, but I have already said more than enough for easy comprehension at one sitting. It is time to pull together what has been accomplished so far and to sight ahead in the direction of its thrust.

We began by noting a major supposition of commonsense confirmation theory, the importance of which lies not in its support (or even recognition) by extant philosophic models of rational inference but in its existentialistic operation as a rule we *live* by. This is that information about the truth of the if-clause in an ordinary-language if/then conditional, or about instances just of the antecedent in an all/are generality, tells nothing relative to natural background information about whether the conjectured dependency itself obtains. Now, until it is normatively clear what propositions *should* be confirmationally relevant/irrelevant to one another under what background information, we can ill afford as philosophers and methodologists of science to disrespect our practical intuitions about this; rather, any discernable consistencies within the latter warrant careful identification and analysis in order that we may be instructed by the principles which govern them. Even if these intuitive principles prove normatively defective in some respects, they may well nonetheless provide (as common sense so often does) an essential first-approximation to what needs be said by a more technically adequate account of the matter, without whose inspiration the latter would never get off the mark at all.

My primary objective in this essay, then, has been to see what can be made of this commonsensical confirmatory indifference of subjunctive conditionals and conditional generalities to data on their antecedents; specifically, to inquire whether it supposes anything of philosophical significance about thought or reality, whether in the large it is epistemologically defensible, and, if its naive version fails to achieve cognitive coherence, whether there may not be *something* in it which technical reconstructions of rational inference need to preserve. This question's philosophic importance is all the greater in that the ordinary-language sense of conditionality at issue here is the one wherein 'If p then q' has roughly the force of saying that q would be *due to* p were the latter the case, that p would *bring about*, or *cause*, or be *nomically responsible* for q. These italicized notions have been notoriously refractory to philosophic comprehension, but the possibility now arises that an essential part of their meaning may reside in or at least become accessible through their behavior in confirmational relationships.

My first excursion into this territory (Rozeboom, 1968) achieved, I think, a useful sighting on heretofore unsuspected wilds of confirmation-theoretic *problems*, but retrieved little in the way of conclusions. Now, however, we have arrived at two results solid enough to build upon. The first is that, common sense to the contrary notwithstanding, if-p-then-q is *not* in general confirmationally independent of p, nor is all-As-are-Bs of Aa, under natural or impoverished background information in *any* sense of conditionality. Hence in particular, this cannot be an intrinsic credibility property of nomic conditionals. Even so—and this is our second main result, though I have not yet quite finished putting it together— there is a statistical counterpart to this commonsense coupling of causal-order suppositions with our uncertainty structure which not only appears logically coherent when suitably qualified but which captures remarkably well the spirit of the original intuition albeit at the price of greater technical complexity. This is that if something's being or not being an A is assumed to have complete causal precedence over its being or not being a B— i.e. given that any causal connection between these variables goes exclusively from A to B—while our remaining background information is sufficiently natural, then our uncertainty about $\text{fr}(A)$ is independent of our uncertainty about $\text{fr}(B \mid A)$, with the result that whether or not a particular object a is an A is irrelevant to whether or not statistically all As are Bs. It is worth making clear just how this statistical revision of the nomic indifference principle improves upon the original. In both cases, the driving idea is that our construing one event, or one kind of event, to be responsible for another is mirrored somehow by confirmational independence between such conditionalities and their antecedents. In the original intuition, however, this nomic force is projected into the if/then and all/are content of the statements to which the corresponding if-antecedents are then thought irrelevant. In contrast, since statements about conditional frequencies are neutral regarding which properties have responsibility for bringing about which others, the intuitive principle's statistical emendation withholds nomic force from the conditionality expressed by 'All As are Bs' and its implicate 'If Aa then Ba' but puts it into the background under which the latter are deemed confirmationally independent of Aa.

Where do we go from here? I'm sure that you are sceptical that principle *CSC* can survive much hostile examination, and I must confess to my own qualms about this. Yet there is a right way and a wrong way to go about criticizing *CSC*. The wrong one would be to ignore it altogether as philosophically outlandish (i.e. disquietingly unorthodox) or dismiss it on grounds that your own causal concepts just don't work that way. Such an argument, however, would presume that your present intuitive grasp of causation is so unproblematically clear that any proposed explication of nomic responsibility which disagrees with it must perforce be misguided. If the whole sorry history of philosophic fumbling with causal concepts proves anything at all, it is surely that we do *not* understand these at all well. Certain attitudes of ordinary-language philosophy to the contrary notwithstanding, our everyday working concepts never achieve the heights of clarity, precision, and consistency which ideally they should enjoy, but are under constant pressure to evolve as the individuals and cultures which employ them mature in experience and intellectual sophistication. Accordingly, it might be argued that *CSC* expresses not so much an incontrovertibly analytic property of nomic structure as we now conceive of this as it is a persistent theme therein which offers the best foundation for a technically superior reworking of this still-obscure notion. That is, if some version of *CSC* is not now true of your de facto worldview, perhaps it *ought* to be.

The most auspicious way to appraise *CSC*, it seems to me, is through more general inquiry into the significance of confirmational independence in statistical inference. If it is not rationally arbitrary which parameters of a given distribution family are confirmationally independent given information k—and if this *is* arbitrary the logic of induction is in deep, deep trouble—then it is not at all implausible that whatever k implies about causal order has major bearing on this. If so, then deciphering the nature of that import will in passing also critique *CSC*, which, meanwhile, will usefully serve to model how information about nomic structure *might* control the structure of statistical credibilities. However, *CSC* by no means delimns the breadth or horizons of this front of inquiry. Ultimately at issue is a much more basic possibility: When how we manage our beliefs has pragmatic repercussions foreseeably serious enough to evoke anxiety whether we are doing this *correctly*, then may it not be that we are inexorably driven to create conceptions of states of affairs which, were they to obtain, would *justify*—i.e. insure the correctness of—our managing our beliefs one way rather than another? The notion of statistical probability may well have had such an origin, namely, as justification for degrees of subjective confidence;[15] the intuitive credibility decoupling of law-like conditionals from their antecedents urges that our still-formative ideas about nomic responsibility have a similar basis; and it is not too much to hope that still others of the modal connectives and operators which have for so long frustrated analytic philosophers will reveal hitherto unsuspected transparencies when viewed from this perspective.

Meanwhile, the prospect directly at hand is that not only do causal suppositions determine the structure of our uncertainties *if* we happen to believe in such (as viewed by some) philosophically dubious event-couplings, but conversely that without some such commitment we have no rational basis on which to allocate confirmational dependencies—specifically, that whatever can be cited to justify one confirmational structure over another will also serve to define what it is for

[15] Why should I feel rather sure, considering the appearance of the sky, that it will rain tonight? Because when the sky looks like that the chance of rain is about 70%.

one event, or event type, to be nomically responsible for another. How such definitions can best be shaped will depend upon just what logical connections between the two structures, causal and confirmational, prove to be most analytically fundamental; but we may anticipate constructions *roughly* to the effect that if S is the totality of confirmational properties necessarily possessed by a rational system of conditional credibilities given causal hypothesis H, while $\{C_i\}$ is a suitably comprehensive set of conditions less philosophically problematic than H such that each C_i is sufficient reason for a credibility system to have properties S, then H may be explicated as what is common to these C_i.

I conclude, then, with a revitalized hope and an operational directive for the philosophy of causation. The hope arises from discovery that the closed circuit of blind alleys within which this concept's analysis has for so long been pursued opens upon heretofore unknown corridors which reach in promisingly new directions. It would be naive to expect these to lead us directly into the naked light of total comprehension, but at the very least they should merge our currently disparate confusions on several major philosophic issues into a single integrated perplexity which substantially reduces our total uncertainty over these matters even if the marginal uncertainty in any one is not greatly diminished. As for the directive, this is not merely that we search out the confirmational implications of causal hypotheses as we now intuit these, but that especial effort be made to diagnose the conditions under which two propositions approach confirmational independence under real-life background information. For roundabout as this might seem, detailed study of *approximate* irrelevance in statistical inference is perhaps the most powerful research press now at our avail for cracking the causal conundrum. This is because if ξ and Ω are two parameters of some unknown statistical distribution, then just as our uncertainty about ξ or Ω given k_i may converge under a series of empirical data $k_1, k_2, \ldots, k_i, \ldots$ (notably, where k_i describes observed frequencies in a size-i sample of events) to asymptotic conviction in a particular value determined by k_i regardless of our unconditional credibility distribution for this parameter, so may statistical intuition also insist that for some such k-series, $\mathrm{CR}(\xi, \Omega \mid k)$ converges to unity with increasing i. If so, k_i stands revealed as de facto *empirical evidence* for whatever causal conjecture is required to justify treating ξ and Ω as perfectly independent, where the larger is i the more conclusively does k_i support this causal inference. My concluding proposal, in short, is that nomic structure may well relate to the limit of certain data aggregates in much the same way as statistical probabilities relate to limiting relative frequencies. Even if statistical probability cannot convincingly be *identified* with relative frequency, it is nonetheless illuminating to know that these are asymptotically equivalent numerically; and if it can be established that special kinds of empirical evidence similarly converge to causal conclusions, we shall finally have achieved our first significant domestication of this indispensable but still-feral concept.

References

Blalock, H. M. (1961). *Causal inferences in nonexperimental research.* Chapel Hill, NC: The University of North Carolina Press.

Hacking, I. (1967). Slightly more realistic personal probability. *Philosophy of Science*, *34*, 311–325.

Raiffa, H., & Schaiffer, R. (1961). *Applied statistical decision theory.* Boston, Mass.: Harvard University.

Rozeboom, W. W. (1961). Ontological induction and the logical typology of scientific variables. *Philosophy of Science*, *28*, 337-377.

Rozeboom, W. W. (1966). Scaling theory and the nature of measurement. *Synthese*, *16*, 170–233.

Rozeboom, W. W. (1968). New dimensions of confirmation theory. *Philosophy of Science*, *35*, 134–155.

Rozeboom, W. W. (1969). New mysteries for old: The transfiguration of Miller's paradox. *British Journal for Philosophy of Science*, *19*, 345–353.

Whittle, P. (1957). Curve and periodogram smoothing. *Journal of the Royal Statistical Society, Series B*, *19*, 1615–1627.

Whittle, P. (1958). On the smoothing of probability density functions. *Journal of the Royal Statistical Society, Series B*, *20*, 334–343.