

the identification is made. The point is that from a broad perspective the search for causes and the measurement of their effects is a seamless enterprise moving back and forth between closely related questions. The absence of a causal language in much experimentation in the social sciences is due to a positivistic tradition that began very early in this century, if not earlier, and is only slowly dying out.

**Counterfactuals.** Paul properly emphasizes the place where counterfactuals enter and the theory of experiments with treatments and control units begins. What I want to emphasize is that they enter *en masse* whenever we have a probability space. Thus, especially in the theoretical case where the probability space may be huge, if we were really careful in our language when we talk about conditional probabilities we could use counterfactual language. Such and such would have a probability if it were the case that such and such were so. In fact, this use of the subjunctive is gradually fading and is not present now in much scientific discussion of these matters. Although the shoe seems to pinch more in many direct empirical cases, I do not see in principle any difference between such cases and the use of counterfactuals in theoretical contexts, since most of the conditionals introduced will not be realized. What I do think is that no special logic of counterfactuals is needed for either empirical or theoretical purposes. I hope, but I am not quite sure that Paul agrees with me on this point.

#### REFERENCE

Suppes, P.: 1970, *A Probabilistic Theory of Causality*, Amsterdam, North Holland.

#### CAUSAL TENDENCY, NECESSITIVITY AND SUFFICIENTIVITY: AN UPDATED REVIEW

**ABSTRACT.** The tendency of one event  $F$  to cause a later event  $E$  was explicated by a simple formula in terms of probability by Good (1961/62), more specifically in terms of 'weights of evidence' (logarithms of Bayes factors). The argument was based on the concept of causal circuits. The same explication, and a generalization, are reached here without making use of that concept, and the mathematical level is elementary.

It is recognized here that the explication would be more explicitly described as capturing the notion of the tendency of  $F$  to be a *sufficient* cause of  $E$ , and a similar explication is proposed for the tendency to be a *necessary* cause. For convenience in discussing these explications they are called *sufficientity* and *necessitivity*. They bear a simple dual or skew-symmetric relationship to each other and have simple additive properties.

For supplementary material see especially my contribution to Jarvie (forthcoming), as well as the monograph by Suppes (1970).

I have recently written (Good, 1992a) a brief review of my work on the mathematics of philosophy which included work on causal tendency. Causal tendency has greatly interested both Patrick Suppes (1970) and myself for decades, so I felt it was appropriate to return to the topic in the present volume, and to add something to what I said in Good (1992a) concerning the quantitative concept of the 'tendency to be a sufficient (or necessary) cause'. That paper may be regarded as 'supplementary reading'. Some qualitative implications will also be mentioned. My aim is to make the theory seem as natural as, say, Newton's three laws of motion. Perhaps my theory will break down, too, when it is necessary to allow for the weirdities of quantum theory.

There are many formulae in this paper but the mathematics is elementary, roughly at the level of a simple discussion of Boolean algebra. So 'non-mathematical' philosophers of science should not be alarmed. In the philosophy of science, a little mathematics is often necessary and sufficient.

A measure of causal tendency should be useful in the theory of neural circuits, whether artificial or natural. For my speculations on this

point see Good (1961, p. 228; 1965; and implicitly in 1988a, p. 25). The concept of causal circuits was a main theme of Good (1961/62 and 1988a) and was used to arrive at a measure for causal tendency, but my justification for this same measure in the present work is quite different. The basic idea is to 'probabilize', in the simplest possible way, the concepts of strictly sufficient and strictly necessary causality. A potential application would be for improving your body of beliefs by adding to your variety of judgments (of probabilities, odds, Bayes factors, weights of evidence, utilities and expected utilities). Several decades might pass before such applications become routine.

The distinction between sufficient causation and necessary causation is important in the philosophy of nominally 'strict' (deterministic) causality. For example, in the index of Hart and Honoré (1959), there are 45 page references to 'necessary conditions', under the heading *conditio sine qua non*, and ten to 'sufficient conditions'. See also Salmon (1984, pp. 185–189). Of course these authors do recognize the importance of probabilistic considerations. Strict causality should be a limiting case of probabilistic causality, a principle that acts as a check of the theory. Suppes (1970, p. 34) calls a strict cause a *sufficient* cause and, on his p. 75, attributes this description to Mill and others.

Until recently, when discussing probabilistic causality, I overlooked the need to distinguish<sup>1</sup> between the degree to which an event  $F$  tends to be a sufficient cause and, on the other hand, the tendency to be a necessary cause of a later event  $E$ , later in the sense that  $F$  or its negation (whichever occurs) ends not later than  $E$  or its negation begins. Perhaps I was counter-influenced by the lack of convenient terminology,<sup>2</sup> somewhat as we would be handicapped if words like *probability*, *probably*, and *chance* didn't exist. Accordingly I think it is advisable to begin with a discussion of terminology and notation. Instead of the clumsy expressions 'the degree to which  $F$  tends to be a necessary (or sufficient) cause of  $E$ ', I propose the shorter expressions:<sup>3</sup>

- (1) the necessitivity of  $F$  for  $E$  (given  $U$ )

or

- (2) the necessitatedness of  $E$  provided by  $F$  (given  $U$ ),

and

- (3) the sufficientivity of  $F$  for  $E$  (given  $U$ )

or

- (4) the sufficientatedness of  $E$  provided by  $F$  (given  $U$ ),

where  $U$  is defined in Note 3. The corresponding symbolic notations, which so far contain no mathematics, are

- (5)  $Q_{\text{nec}}(E : F | U)$  and  $Q_{\text{suf}}(E : F | U)$ .

Of the four expressions, (1) and (3) are the more euphonic pair, whereas (2) and (4) enable one to read the notations from left to right, with the colon read as *provided by*. I mention this less euphonic pair mainly to justify the use of the colon here as used elsewhere in relation to weights of evidence and to Bayes factors.

Note that I am assuming that the colon *takes precedence* over, or is *more binding* than the vertical stroke in the sense that it is unnecessary to enclose  $E : F$  in parentheses. The idea of a precedence order for notations in logic, to economize in the use of parentheses and thereby to increase readability, is used, for example, by Hilbert and Ackermann (1950, pp. 7 and 66), just as multiplication takes precedence over addition in elementary algebra (a precedence that is emphasized by the frequent omission of the multiplication symbol, or its replacement by a dot).

The complete precedence rules that I adopt are: negation takes precedence over conjunction, which takes precedence over disjunction, which takes precedence over the oblique stroke (explained later), which takes precedence over the colon, which takes precedence over the vertical stroke. Now let's go back to 'square one'.

An *event* is something that might occur or might not. Our events will be of bounded spacial and temporal size. Very often, in ordinary usage, an event is regarded as a bundle or disjunction of many more specific events that might be called *microevents*. This matter is discussed by Good (1988a,b) while Suppes (1970) gives more detailed discussions of the meaning of an 'event', – see the index of his book.

A *proposition* is the meaning of a statement. Some symbols for propositions will be  $A, B, C, D, E, F, G, H$ , sometimes with subscripts or other 'qualifications'. In particular,  $E$  will usually correspond to an event, experimental results, or evidence. Indeed,  $E$  might ambiguously denote both an event and the proposition that the event  $E$  'obtains'. The symbol  $H$  often denotes a hypothesis or theory. The negation<sup>4</sup> of  $E$ , for example, is denoted by  $\bar{E}$ . The conjunction of, say,  $A$  and  $B$  is denoted

by  $A \& B$  or by  $A.B$  (the period never being omitted herein) while the disjunction is denoted by  $A \vee B$  (true if  $A$  or  $B$  or both are true). De Morgan's law is  $\overline{A \& B} = \overline{A} \vee \overline{B}$ . Of course  $\overline{A \& B}$  is not the same as  $\overline{A} \& \overline{B}$  (just as most educated people, when they are not tired or drunk, distinguish between 'not all people are' and 'all people are not').

$P(A | B)$  (with a vertical stroke) denotes, as is customary in mathematical texts, the probability of  $A$  given  $B$ , where, in this article, for the sake of the soul of wit, it will be left to the reader to decide whether a probability is: (i) a propensity, that is, a physical rather than an epistemic probability (a logical probability or a personal probability); or (ii) that both interpretations are acceptable. For my opinions concerning kinds of probability see, for example, Good (1983). For the sake of simplicity I shall regard all probabilities as having sharp values although, as I have emphasized on dozens of occasions, I think interval-valued probabilities are more realistic, at least when they are subjective (personal) probabilities. I believe that everything in this chapter could be rewritten in terms of interval-valued probabilities but that the trees would then obscure the forest. To put the matter epigrammatically, it is better to communicate briefly and partially than fully and not at all.

Odds,  $O(A | B)$ , means  $P(A | B)/P(\overline{A} | B)$ , but  $A/B$  means  $A$  as compared with or as contrasted with or as against  $B$ . (Of course the oblique stroke between numbers denotes division.) For example,  $O(A/B | C)$  means  $P(A | C)/P(B | C)$ . Thus  $O(A/\overline{A} | C)$  is another way of writing  $O(A | C)$ . We can think of  $O(A/B | C)$  as a generalization of  $O(A | C)$  and can read it as the odds of  $A$  relative to or as against  $B$  (given  $C$ ). The reader should stop to check that  $O(A/B | C)$  is not equal to  $O(A | C)/O(B | C)$  unless  $P(A | C) = P(B | C)$ .

The factor by which the prior odds of a hypothesis  $H$  are multiplied, in the light of evidence  $E$ , given background information  $G$ , to obtain the posterior odds of  $H$ , is called the Bayes factor in favor of  $H$  provided by  $E$  given  $G$ , and it may be denoted by  $BF(H : E | G)$  with no dot between  $B$  and  $F$ ! The colon, pronounced *provided by* in this context, should not be confused with the vertical stroke, pronounced and meaning *given*.

The extremely important odds form of Bayes's theorem (a form that Bayes never mentioned, but Poisson came close) is

$$(6) \quad BF(H : E | G) = \frac{O(H | E.G)}{O(H | G)} = \frac{P(E | H.G)}{P(E | \overline{H}.G)},$$

and any reader unfamiliar with it should write out the proof.<sup>5</sup> We always assume at least implicitly that ratios are not of the indeterminate form 0/0. Thus the Bayes factor couldn't be calculated if one were to observe an event that was impossible or 'almost impossible' (such as a 'continuous' observation made with infinite accuracy) given  $H.G$  and also given  $\overline{H}.G$ . If such an event occurred then  $G$  would be refuted!

A useful generalization of the Bayes factor is

$$(7) \quad BF(H_1/H_2 : E | G) = \frac{O(H_1/H_2 | E.G)}{O(H_1/H_2 | G)} \\ = \frac{P(E | H_1.G)}{P(E | H_2.G)}$$

(where the left equation is a definition), and the proof of this is also a simple exercise. The right side is sometimes meaningful to a person who thinks he is a non-Bayesian, namely when it is a simple 'likelihood ratio'. In that case, Equation (7) gives an intuitive interpretation to the likelihood ratio.

The odds form of Bayes's theorem shows that any entirely reasonable definition of the weight of evidence  $W$  in favor of a hypothesis, provided by  $E$ , given  $G$ , must be a strictly increasing function of the Bayes factor. This claim can also be shown (for example, by Good, 1989) by a brief argument, now repeated, that is sometimes non-Bayesian.<sup>6</sup> We assume the desiderata (i) that  $W(H : E | G)$  depends only on  $x = P(E | H.G)$  and on  $y = P(E | \overline{H}.G)$ , say

$$(8) \quad W(H : E | G) = f(x, y, G) = f(x, y)$$

by definition, if  $G$  is fixed; and (ii) that  $f$  is an increasing function of  $x$  for any fixed value of  $y$  (and  $G$ ). Both desiderata seem compelling to me. Suppose now that  $E^*$  has nothing to do with  $E$  or  $H$ , given  $G$ . Then

$$(9) \quad P(E.E^* | H.G) = P(E^* | G)P(E | H.G) = \lambda x,$$

where  $\lambda = P(E^* | G)$ , and

$$(10) \quad P(E.E^* | \overline{H}.G) = P(E^* | G)P(E | \overline{H}.G) = \lambda y.$$

But the equation  $W(H : E.E^* | G) = W(H : E | G)$  is intuitively obvious, and therefore  $f(\lambda x, \lambda y) = f(x, y)$ . But  $\lambda$  could take any

positive value between 0 and 1, by appropriate choice of  $E^*$ , and it follows that  $f(x, y)$  is some function of  $x/y$ , that is, of the Bayes factor, or of the simple likelihood ratio under the non-Bayesian interpretation. The claim now follows from the second 'compelling' desideratum.

The argument generalizes readily to show that  $W(H_1/H_2 : E | G)$  must be an increasing function of the generalized Bayes factor  $BF(H_1/H_2 : E | G)$ . To obtain the additive property

$$(11) \quad W(H_1/H_2 : E.F | G) \\ = W(H_1/H_2 : E | G) + W(H_1/H_2 : F | E.G)$$

we must define weight of evidence as the logarithm of the Bayes factor. The base of the logarithms can be any number exceeding unity and it merely determines the unit in terms of which weights of evidence are measured. The base is the same, but unspecified, throughout this chapter; for example, it might be 10 or  $10^{0.1}$  or  $e$  or 2, in which cases the units are respectively bans, or decibans, or natural bans, or bits. The names *ban*, *deciban* and *natural ban* were suggested by Turing (1941). See also Good (1992c).

Although additivity is less important than the other desiderata it is a simple and convenient property. Thus it leads to the best explication of the concept of weight of evidence. In the case where  $E$  and  $F$  are statistically independent given  $H_1.G$  and also given  $H_2.G$ , we obtain the simpler additive property

$$(12) \quad W(H_1/H_2 : E.F | G) \\ = W(H_1/H_2 : E | G) + W(H_1/H_2 : F | G)$$

which is needed if Themis's scales are taken seriously, that is, if the ancient Grecian concept of weight of evidence is to be thoroughly captured. Of course, by taking antilogarithms of (11) and (12) we obtain the corresponding multiplicative properties of Bayes factors.

Anyone who still doesn't agree that  $W$  is the best explication of weight of evidence must be a Marxist (or Grouchemarxist) in the sense of the *great* Marx (1932) who said "Whatever it is, I'm against it".<sup>7</sup>

A useful property of Bayes factors and weights of evidence is that they are barely affected by, or are 'robust with respect to', a description of  $E$  with unnecessary detail, for example, to more places of decimals than necessary or as a microevent. This property is closely related to the property  $W(H : E.E^* | G) = W(H : E | G)$ , which we described above as 'intuitively obvious' (if 'indeterminate forms' are excluded).

I shall now argue that the best explications for sufficientity and necessity are respectively (indeterminate forms being excluded):

$$(13) \quad Q_{\text{suf}}(E : F | U) = W(\bar{F} : \bar{E} | U)$$

and

$$(14) \quad Q_{\text{nec}}(E : F | U) = W(F : E | U).$$

(If either of these expressions is negative then so is the other and then  $F$  is a *negative cause* of  $E$ . Without the notation  $U$  we have what Suppes (1970, pp. 43 and 44) calls "prima facie negative causes".) More generally, I claim that the best explication, for the sufficientity for  $E$ , of  $F_1$  as against  $F_2$  (if  $P(\bar{E} | F_1.U)$  and  $P(\bar{E} | F_2.U)$  are not both zero) is

$$(15) \quad Q_{\text{suf}}(E : F_1/F_2 | U) = W(F_2/F_1 : \bar{E} | U),$$

where, in the notation,  $\bar{F}$  has become  $F_2/F_1$ , and that for the necessity the best explication (if  $P(E | F_1.U)$  and  $P(E | F_2.U)$  are not both zero) is

$$(16) \quad Q_{\text{nec}}(E : F_1/F_2 | U) = W(F_1/F_2 : E | U).$$

The explication (13) is the same as the one reached in Good (1961/62) by arguing in terms of causal circuits, except that I did not then use the subscript  $\text{suf}$ . The following arguments are similar to, but not identical to, a part of Good (1992a) and are more general because they justify (15) and (16) instead of the special cases (13) and (14).

Before presenting the main arguments let's consider the special cases of strict sufficiency and strict necessity. Suppose first that  $F$  is strictly sufficient for  $E$ . Then if  $E$  doesn't occur,  $F$  couldn't have occurred. Second, suppose that  $F$  is strictly necessary for  $E$ . Then if  $E$  occurs,  $F$  must have occurred. Thus, in the cases of strict sufficiency and strict necessity, the right sides of (13) and (14) are respectively infinite. This gives some initial support for our explications. We now proceed with the main arguments.

We begin with the natural desideratum that  $Q_{\text{suf}}(E : F_1/F_2 | U)$  and  $Q_{\text{nec}}(E : F_1/F_2 | U)$  are both functions of  $\xi_1 = P(E | F_1.U)$  and  $\xi_2 = P(E | F_2.U)$ . It will be convenient to write  $\eta_1$  for  $1 - \xi_1$  and

$\eta_2$  for  $1 - \xi_2$ . Note that of course any function of  $\xi_1$  and  $\xi_2$  is also a function of  $\eta_1$  and  $\eta_2$ . We adopt the notations  $g$  and  $h$  where

$$(17) \quad Q_{\text{suf}}(E : F_1/F_2 | U) = g[P(\bar{E} | F_1.U), P(\bar{E} | F_2.U)] \\ = g(\eta_1, \eta_2)$$

and

$$(18) \quad Q_{\text{nec}}(E : F_1/F_2 | U) = h[P(E | F_1.U), P(E | F_2.U)] \\ = h(\xi_1, \xi_2).$$

Let  $E^*$  (neither 'almost certain' nor 'almost impossible,' to avoid indeterminate forms<sup>8</sup>) denote some event that is entirely independent of  $E$ ,  $F_1$ , and  $F_2$  (given  $U$ ). Assume, as further desiderata,

$$(19) \quad (i) \quad Q_{\text{suf}}(E \vee E^* : F_1/F_2 | U) = Q_{\text{suf}}(E : F_1/F_2 | U)$$

and

$$(20) \quad (ii) \quad Q_{\text{nec}}(E.E^* : F_1/F_2 | U) = Q_{\text{nec}}(E : F_1/F_2 | U)$$

because both desiderata are clearly true in the special case of strict causality and intuitively highly appealing when the causality is not strict. (We shall see later that Equations (19) and (20) are special cases of additive formulae, namely (37) and (36).)

A further desideratum, which is entirely compelling, is

(iii) for any fixed  $U$  and fixed value of  $P(E | F_2.U)$ , both  $Q_{\text{suf}}(E : F_1/F_2 | U)$  and  $Q_{\text{nec}}(E : F_1/F_2 | U)$  are increasing functions of  $P(E | F_1.U)$ .

On substituting  $E \vee E^*$  for  $E$  in the identity (17), we obtain

$$(21) \quad Q_{\text{suf}}(E \vee E^* : F_1/F_2 | U) \\ = g[P(\overline{E \vee E^*} | F_1.U), P(\overline{E \vee E^*} | F_2.U)] \\ = g[P(\bar{E}.\bar{E}^* | F_1.U), P(\bar{E}.\bar{E}^* | F_2.U)] \\ \text{(by de Morgan's law)} \\ = g[P(\bar{E}^* | U)P(\bar{E} | F_1.U), P(\bar{E}^* | U)P(\bar{E} | F_2.U)] \\ \text{(from the definition of } E^*) \\ = g(\mu\eta_1, \mu\eta_2)$$

where  $\mu = P(\bar{E}^* | U)$ . But, by our Desideratum (i) or (19), the left side of the identities (21) is equal to

$$(22) \quad Q_{\text{suf}}(E : F_1/F_2 | U) = g(\eta_1, \eta_2)$$

and therefore

$$(23) \quad g(\eta_1, \eta_2) = g(\mu\eta_1, \mu\eta_2)$$

is an identity, and therefore  $g(\eta_1, \eta_2)$  is a function of the ratio  $\eta_1/\eta_2$ . Equivalently it is a function of  $W(F_1/F_2 : \bar{E})$ . But this is equal to  $-W(F_2/F_1 : \bar{E})$  (because the logarithm of  $a/b$  is minus the logarithm of  $b/a$ ) so  $Q_{\text{suf}}(E : F_1/F_2 | U)$  is a function of  $W(F_2/F_1 : \bar{E})$ . It must be an increasing function to satisfy Desideratum (iii).

We may as well take the simplest function,<sup>9</sup> as in our claim (15). This has the further advantage of leading to appealing additive properties that are discussed below. This completes the argument for claim (15).

*Exercise.* Prove and generalize the identity

$$(13A) \quad Q_{\text{suf}}(E : F | U) = \log \left[ \frac{1 + O(E | F.U)}{1 + O(E | \bar{F}.U)} \right].$$

The argument for claim (16) resembles that for (15) but is slightly simpler. I give it here for the sake of completeness.

On substituting  $E.E^*$  for  $E$  in the identity (18) we have

$$(24) \quad Q_{\text{nec}}(E.E^* : F_1/F_2 | U) \\ = h[P(E.E^* | F_1.U), P(E.E^* | F_2.U)] \\ = h[P(E^* | U)P(E | F_1.U), P(E^* | U)P(E | F_2.U)] \\ = h(\lambda\xi_1, \lambda\xi_2)$$

where  $\lambda = P(E^* | U)$ . But, by Desideratum (ii) or (20), the left side of the identities (24) is equal to

$$(25) \quad Q_{\text{nec}}(E : F_1/F_2 | U) = h(\xi_1, \xi_2)$$

and therefore

$$(26) \quad h(\lambda\xi_1, \lambda\xi_2) = h(\xi_1, \xi_2)$$

is an identity and therefore  $h(\xi_1, \xi_2)$  is a function of the ratio  $\xi_1/\xi_2$ . Equivalently it is a function of  $W(F_1/F_2 : E)$ . Therefore, by Desideratum (iii), it must be an increasing function. We may as well take the simplest function as in our claim (16). This again leads to additive properties.

**COROLLARY OF CLAIMS (15) AND (16).** *We have the following fundamental identity, or generalized or probabilistic contraposition, or principle of duality or asymmetry or antisymmetry or skew-symmetry:*

$$(27) \quad Q_{\text{nec}}(E : F_1/F_2 | U) = Q_{\text{suf}}(\bar{E} : F_2/F_1 | U),$$

*a special case of which is*

$$(28) \quad Q_{\text{nec}}(E : F | U) = Q_{\text{suf}}(\bar{E} : \bar{F} | U).$$

Formula (28) expresses a probabilistic form of the familiar logical principle of contraposition. It was called a *fundamental identity* by Good (1992a) (and the same description is appropriate for (27)), where the example was given that going out of doors is strongly necessary for getting run over (killed by being hit by a vehicle), whereas staying at home is strongly sufficient for not getting run over. This of course meant that the corresponding necessity and sufficiency are large, but they are not infinite because traffic and aircraft do sometimes crash into houses. This corollary has strong intuitive appeal and so provides further support to our explications.

We express (28) in words: to pass from necessity to sufficiency, or vice versa, just negate both 'arguments' that precede the vertical stroke. Likewise, to express (27) in words, negate the first 'argument' and interchange the subscripts 1 and 2 in the second 'argument'.

#### ADDITIVITY

The following additive property of  $Q_{\text{suf}}$  is mentioned by Good (1961/62, p. 44), though with a different notation (especially in that I had not yet seen the merit of the subscript  $_{\text{suf}}$ ).

$$(29) \quad Q_{\text{suf}}(E : F_1.F_2/\bar{F}_1.\bar{F}_2 | U) \\ = Q_{\text{suf}}(E : F_1 | F_2.U) + Q_{\text{suf}}(E : F_2 | \bar{F}_1.U).$$

Because the left side of (29) is symmetric with respect to the subscripts 1 and 2, we can interchange these subscripts on the right. This gives

$$(29A) \quad Q_{\text{suf}}(E : F_1.F_2/\bar{F}_1.\bar{F}_2 | U) \\ = Q_{\text{suf}}(E : F_2 | F_1.U) + Q_{\text{suf}}(E : F_1 | \bar{F}_2.U).$$

In our present terminology we can express (29) as follows:

For  $E$ , the sufficientivity of  $F_1.F_2$  as against  $\bar{F}_1.\bar{F}_2$  is equal to that of  $F_1$  given  $F_2$  plus that of  $F_2$  given  $\bar{F}_1$ .

For example, suppose you are compelled to choose between two jobs; in one of them the air is smoky and you are allowed to smoke (and would), whereas, in the other job the air is clean and smoking is disallowed. Then, *ceteris paribus*, the sufficientivity of the first choice, for getting lung cancer, is equal to the sum of the sufficientivities from smoking, *given* that the air is smoky on the one hand, and, on the other hand, from living in a smoky atmosphere *given* that you don't smoke. To make the argument more rigorous we can regard the degrees and kinds of both smokiness and smoking as held constant.

I omitted the proof of (29) in my 1961/62 paper because the editor was anxious to save space, but I give it now to be reader-friendly and take  $U$  for granted in the proof. The right side of (29) equals

$$(30) \quad W(\bar{F}_1 : \bar{E} | F_2) + W(\bar{F}_2 : \bar{E} | \bar{F}_1) \\ = \log \frac{P(\bar{E} | \bar{F}_1.F_2)}{P(\bar{E} | F_1.F_2)} + \log \frac{P(\bar{E} | \bar{F}_1.\bar{F}_2)}{P(\bar{E} | \bar{F}_1.F_2)} \\ = \log \frac{P(\bar{E} | \bar{F}_1.\bar{F}_2)}{P(\bar{E} | F_1.F_2)} \\ [\text{because } \log(a/b) + \log(c/a) = \log(c/b)] \\ = W(\bar{F}_1.\bar{F}_2/F_1.F_2 : \bar{E}) = Q_{\text{suf}}(E : F_1.F_2/\bar{F}_1.\bar{F}_2)$$

and this is the *left* side of (29) when the symbol  $U$  is restored. Thus (29) is proved.

By using the fundamental identities (27) and (28), we deduce from (29A) that

$$(31) \quad Q_{\text{nec}}(\bar{E} : \bar{F}_1.\bar{F}_2/F_1.F_2 | U) \\ = Q_{\text{nec}}(\bar{E} : \bar{F}_1 | \bar{F}_2.U) + Q_{\text{nec}}(\bar{E} : \bar{F}_2 | F_1.U).$$

(I have interchanged the two terms on the right.) But, by a change of notation, namely by replacing  $E$ ,  $F_1$  and  $F_2$  by their negations, the identity (31) becomes

$$(32) \quad Q_{nec}(E : F_1.F_2/\bar{F}_1.\bar{F}_2 | U) \\ = Q_{nec}(E : F_1 | F_2.U) + Q_{nec}(E : F_2 | \bar{F}_1.U)$$

which is exactly of the same form as (29) but with  $_{suf}$  changed to  $_{nec}$ . That is,  $Q_{suf}$  and  $Q_{nec}$  satisfy a *same* form of additivity. Because this might seem somewhat surprising, or unconvincing, here is a direct proof of (32) along the lines of the proof of (29). Once again I take  $U$  for granted in the proof. The right side of (32) is equal to

$$(33) \quad W(F_1 : E | F_2) + W(F_2 : E | \bar{F}_1) \\ = \log \frac{P(E | F_1.F_2)}{P(E | \bar{F}_1.F_2)} + \log \frac{P(E | \bar{F}_1.F_2)}{P(E | \bar{F}_1.\bar{F}_2)} \\ = \log \frac{P(E | F_1.F_2)}{P(E | \bar{F}_1.\bar{F}_2)} \\ = W(F_1.F_2/\bar{F}_1.\bar{F}_2 : E) = Q_{nec}(E : F_1.F_2/\bar{F}_1.\bar{F}_2),$$

which is the left side of (32).

The following additional additive properties may be proved as exercises.

$$(34) \quad Q_{nec}(E_1.E_2 : F | U) \\ = Q_{nec}(E_1 : F | U) + Q_{nec}(E_2 : F | E_1.U)$$

$$(34A) \quad Q_{nec}(E_1.E_2 \dots E_n : F | U) \\ = \sum_{r=1}^n Q_{nec}(E_r : F | E_1 \dots E_{r-1}.U)$$

(An 'empty product', which occurs when  $r = 1$ , is interpreted as a tautology.)

$$(35) \quad Q_{suf}(E_1 \vee E_2 : F | U) \\ = Q_{suf}(E_1 : F | U) + Q_{suf}(E_2 : F | \bar{E}_1.U)$$

$$(35A) \quad Q_{suf}(E_1 \vee E_2 \vee \dots \vee E_n : F | U) \\ = \sum_{r=1}^n Q_{suf}(E_r : F | \bar{E}_1 \dots \bar{E}_{r-1}.U)$$

More generally than (34) and (35), we have

$$(36) \quad Q_{nec}(E_1.E_2 : F_1/F_2 | U) \\ = Q_{nec}(E_1 : F_1/F_2 | U) + Q_{nec}(E_2 : F_1/F_2 | E_1.U)$$

and

$$(37) \quad Q_{suf}(E_1 \vee E_2 : F_1/F_2 | U) \\ = Q_{suf}(E_1 : F_1/F_2 | U) + Q_{suf}(E_2 : F_1/F_2 | \bar{E}_1.U).$$

The similar generalizations of (34A) and (35A) now go without saying. The identities (36) and (37) are generalizations of the desiderata (20) and (19). Still more generally

$$(38) \quad Q_{nec}(E_1.E_2 : F_1.F_2/\bar{F}_1.\bar{F}_2 | U) \\ = Q_{nec}(E_1 : F_1 | F_2.U) + Q_{nec}(E_1 : F_2 | \bar{F}_1.U) \\ + Q_{nec}(E_2 : F_1 | E_1.F_2.U) \\ + Q_{nec}(E_2 : F_2 | E_1.\bar{F}_1.U)$$

and

$$(39) \quad Q_{suf}(E_1 \vee E_2 : F_1.F_2/\bar{F}_1.\bar{F}_2 | U) \\ = Q_{suf}(E_1 : F_1 | F_2.U) + Q_{suf}(E_1 : F_2 | \bar{F}_1.U) \\ + Q_{suf}(E_2 : F_1 | \bar{E}_1.F_2.U) \\ + Q_{suf}(E_2 : F_2 | \bar{E}_1.\bar{F}_1.U).$$

An example of formula (39) is that the sufficientity of smoking (a definite amount and in a definite manner) and living in a smoky area (smoky to a definite extent) as against not smoking and living in a clean area, for getting either lung cancer or a heart attack (or both), is equal to the sum of four sufficientivities:

- (i) that from smoking, for getting lung cancer, given that you live in a smoky area;

- (ii) that from living in a smoky area, for getting lung cancer, given that you don't smoke;
- (iii) that from smoking, for getting a heart attack, given that you don't get lung cancer and you live in a smoky area; and
- (iv) that from living in a smoky area, for getting a heart attack, given that you don't get lung cancer and you don't smoke.

### MULTIPLICATIVITY

Multiplicative sufficientity and multiplicative necessity can be defined naturally by taking the antilogarithms of  $Q_{\text{suf}}$  and  $Q_{\text{nec}}$ , in other words by using Bayes factors instead of weights of evidence. The multiplicative definitions might be preferable for some purposes. For example, we might want to say that some argument 'multiplies our estimate of the multiplicative sufficientity by about 20', without any implicit or explicit reference to logarithms. Appropriate names for these concepts would be *sufficientive factor* and *necessitive factor*, somewhat analogous to 'Bayes factor', and appropriate notations would be

$$(40) \quad Q_{\text{nec}}^{\text{mult}}(E : F | U) = \frac{P(E | F.U)}{P(E | \bar{F}.U)},$$

and

$$(41) \quad Q_{\text{suf}}^{\text{mult}}(E : F | U) = \frac{P(\bar{E} | \bar{F}.U)}{P(\bar{E} | F.U)}.$$

### QUALITATIVE DEDUCTIONS

There might be some readers who are not yet convinced by the quantitative explications of  $Q_{\text{nec}}$  and  $Q_{\text{suf}}$  but who will find that everything in Table I appeals to their intuitions. In any case, the table will help to bridge the gap between our quantitative exposition and qualitative thinking (cf. Suppes, 1970, Ch. 3). In this table I write  $p = P(E | F.U)$ ,  $q = P(E | \bar{F}.U)$ , and abbreviate  $Q_{\text{nec}}(E : F | U)$  and  $Q_{\text{suf}}(E : F | U)$  to  $Q_{\text{nec}}$  and  $Q_{\text{suf}}$  in the body of the table. The table refers to the explications (13) and (14), not to (15) and (16). So we have

$$(42) \quad Q_{\text{nec}} = W(F : E | U) = \log(p/q)$$

TABLE I

A qualitative comparison of  $Q_{\text{nec}}(E : F | U)$  and  $Q_{\text{suf}}(E : F | U)$ .

If	Then	
$p = q$	$Q_{\text{nec}} = Q_{\text{suf}} = 0$	(T1)
$p$ and $q$ close together but not close to 0 or 1	$Q_{\text{nec}}$ and $Q_{\text{suf}}$ are small (in absolute value)	(T2)
$p + q = 1$	$Q_{\text{nec}} = Q_{\text{suf}}$	(T3)
$p$ close to 1, and $p + q = 1$	$Q_{\text{nec}} = Q_{\text{suf}}$ is large	(T4)
$p = 1, q \neq 1$	$Q_{\text{nec}} = \infty$ (cf. Suppes, 1970, p. 34)	(T5)
$p > q$	$Q_{\text{nec}}$ and $Q_{\text{suf}}$ are both positive	(T6)
$p < q$	$Q_{\text{nec}}$ and $Q_{\text{suf}}$ are both negative	(T7)
$p$ and $q$ interchanged	The absolute magnitudes of $Q_{\text{nec}}$ and $Q_{\text{suf}}$ are unchanged but their signs are changed	(T8)
$q$ close to 1, but $p$ is much closer to 1	$Q_{\text{nec}}$ is small but $Q_{\text{suf}}$ is large	(T9)
$p$ is small, but $q$ is much smaller	$Q_{\text{nec}}$ is large but $Q_{\text{suf}}$ is small	(T10)

and

$$(43) \quad Q_{\text{suf}} = W(\bar{F} : \bar{E} | U) = \log\left(\frac{1-q}{1-p}\right).$$

Exercise:  $Q_{\text{nec}} = Q_{\text{suf}}$  iff<sup>10</sup>  $p = q$  or  $p + q = 1$ .

### COMMENTS CONCERNING TABLE I

Because the numerical magnitudes of  $Q_{\text{nec}}$  and  $Q_{\text{suf}}$  depend on the base of logarithms, when we say that either is small (or large) we mean that  $Q_{\text{nec}}^{\text{mult}}$  or  $Q_{\text{suf}}^{\text{mult}}$  is close to 1 (or large). If this be understood then the meanings of small and large still remain a matter of taste. I think we could reasonably say, of the multiplicative forms, that 10 or more is large.



One example of (T10) is where  $F$  = going for a walk and  $E$  = getting run over, an example considered earlier. Although  $Q_{nec}$  is large, we go for walks partly because  $Q_{suf}$  is small. A similar example, but quantitatively different, is where  $F$  = smoking cigarettes and  $E$  = getting lung cancer.

Re (T7), negative causation is discussed, for example, by Suppes (1970, p. 43).

$Q_{nec}$  and  $Q_{suf}$  always have the same sign or both vanish, by (T1), (T6), and (T7).

It is possible for  $Q_{nec}$  and  $Q_{suf}$  to be both small or both large or one small and the other large. Indeed, every pair of real values of  $Q_{nec}$  and  $Q_{suf}$  is possible provided only that they are of the same sign or both are zero. This assertion can be readily proved by the following argument.

Suppose  $Q_{nec}$  and  $Q_{suf}$  are both positive. Then  $p > q$ , and  $n > 1$  and  $s > 1$  where  $p/q = n$  and  $(1 - q)/(1 - p) = s$ . We have  $p = nq$  and  $1 - q = s(1 - p)$  and therefore

$$(44) \quad p = \frac{ns - n}{ns - 1} \quad \text{and} \quad q = \frac{s - 1}{ns - 1}.$$

Thus  $0 < q < p < 1$ , so every pair of positive values of  $Q_{nec}$  and  $Q_{suf}$  leads to a possible pair of values of  $p$  and  $q$ . The proof for the case where  $Q_{nec}$  and  $Q_{suf}$  are both negative is essentially the same, while the case where they are both zero is covered by property T1.

If  $E$  is replaced by a microdescription of  $E$ , for example, by measurements to many decimal places,<sup>11</sup> the effect on the numerical value of  $p/q$  might be negligible. In this sense  $Q_{nec}$  might well be 'robust' with respect to microdescriptions of  $E$ . This thought might provoke us to think that  $Q_{suf}$  too should be robust, in the same sense, but, with our explication of  $Q_{suf}$  as  $\log[(1 - q)/(1 - p)]$ , it clearly is not. Rather than undermining our explications, this argument supports them somewhat. The point might become clearer if we consider an example.

Suppose that Professor Moriarty fires a gun at Sherlock Holmes (event  $F$ ), and we have some measure of the degree to which  $F$  was necessary for the death of Holmes soon thereafter (event  $E$ ). It makes intuitive sense to say that the shooting would be at least as necessary were  $E$  described in absurd detail. But this would obviously not be true if the word 'necessary' were replaced by 'sufficient'. Indeed, the degree of sufficiency *should* be regarded as extremely small if the description were extremely detailed. The fact that our explications of necessity

and sufficientity have these properties is then consistent with intuition. For ordinary applications of the concept of a somewhat sufficient cause it is therefore essential to think of  $E$  as a *bundle* (logical disjunction) of numerous microevents rather than as a single microevent. Compare the fact that an event is usually described as a (measurable) *set* in the 'measure theory' of probability.

## SUMMARY AND CONCLUSIONS

Explications have been given for the tendency of an event  $F$  to be a sufficient (or necessary) cause of  $E$ , using the neologisms *sufficientity* and *necessitivity* and the notations  $Q_{suf}(E : F | U)$  and  $Q_{nec}(E : F | U)$ . The explications are, respectively: (i) the weight of evidence against  $F$  provided by the nonoccurrence of  $E$ ; and (ii) that in favor of  $F$  given by the occurrence of  $E$ , always given the state of the universe just before  $F$  or its negation occurred. Because of the elegance and simplicity of the associated theory, and because  $Q_{suf}$  agrees with the  $Q$  of my early approach by causal circuits, I believe that these explications are almost entirely convincing not allowing for the possibility of some essentially different interpretation of probabilistic causality.

In accordance with Parkinson's Law, this work has filled the time available for its completion. I didn't have enough time, space, or knowledge to relate the work properly to that of others on probabilistic and strict causality, nor even to my own efforts on  $\chi$  and causal networks.

Department of Statistics,  
Virginia Polytechnic Institute and State University,  
Blacksburg, VA 24061, U.S.A.

## NOTES

<sup>1</sup> The distinction between the tendency  $Q$  of  $F$  to cause  $E$  on the one hand, and the degree  $\chi$  to which  $F$  actually caused  $E$  on a specific occasion on the other, is also of great importance. In the present paper I am not concerned with  $\chi$ . For my earliest and most recent discussions of  $\chi$  see Good (1961/62 and 1992a), and some of the history of that topic can be traced backwards from the latter paper.

<sup>2</sup> Also, a sufficient cause corresponds more to the usual usage of *cause* than does a necessary cause. For example, the degree of punishment for a murder should depend more on the degree to which the culprit's action was sufficient in his opinion, rather

than necessary, to cause the death of the victim. The victim's death is some evidence that the culprit's sufficientity (defined later in the main text) to cause the death was high. If this sufficientity could be well judged directly, then that judgment, rather than the knowledge that the victim died, should be the basis of the degree of punishment. The usual distinction between attempted murder and actual murder provides only rough justice. Precise justice might be impracticable or available only at 'dem Pearly Gates'.

<sup>3</sup> In previous publications since 1961 I have defined  $U$  as the state of the universe just before  $F$  or its negation occurred.<sup>12</sup> This is adequate if  $F$  (and its negation) is regarded as occupying a single point of time (and this can be assumed for most of the present chapter) otherwise  $U$  would be better defined as the set of all influences on  $F$  or on its negation. For the sake of uniformity of notation I shall continue to use the symbol  $U$ . To economize in notation, I shall further regard the symbol  $U$  as covering, in addition, all true laws of nature (which in some of my writings I have denoted by  $H^*$ ).

Of course, when judging a necessity or a sufficientity we can allow only for a small part of the universe. By defining  $U$  as I have done, I have ensured that I haven't omitted anything relevant to physical causation provided that the present is unaffected by the future.

The use of the symbol  $U$  is closely related to the conditioning on events occurring earlier than  $F$  in the definition of a spurious cause by Suppes (1970, p. 21).

<sup>4</sup> The meaning of the negation of a proposition depends on a human choice and it might be either precise or vague. Once this choice is made, I regard  $Q_{\text{sur}}$  and  $Q_{\text{nec}}$  as defined in terms of physical probabilities which, however, can often be judged only subjectively and approximately (see also Good, 1988a, p. 31). From a pragmatic rather than a physically realistic point of view, it is possible to think of all the probabilities as epistemological or even as subjective (personal).

When the notation  $\bar{E}$  is used it is regarded as certain that either  $E$  or  $\bar{E}$  'obtains'. This law of the excluded middle (which is not accepted by 'mathematical intuitionists') makes it unnecessary to write  $E \vee \bar{E}$  to the right of the vertical stroke.

<sup>5</sup> An early mention of (6), but without the terminology (and notation for) 'odds', occurs in Wrinch and Jeffreys (1921) and less clearly in Peirce (1878) (who seemed to assume that the initial odds were 1). I believe that the terminology of odds, Bayes factors, and weights of evidence, has helped considerably in spreading the knowledge of an important concept.

<sup>6</sup> One reason for repeating the brief argument is that it makes music with the arguments for the explications of  $Q_{\text{sur}}$  and  $Q_{\text{nec}}$ , (15) and (16).

<sup>7</sup> The expression 'weight of evidence' is used *informally*, for example, by Fisher (1938, pp. 83 and 84; 1956, pp. 98 and 100) and by Kempthorne and Doerfler (1969, p. 233) as the interpretation of a tail-area probability or  $P$ -value. Their appeal is clearly to the reader's intuitive and vague concept of weight of evidence and not to an *explication* of this concept. For some further discussion of, and references concerning, this matter see, for example, Good (1991 or 1992b) where it is emphasized that a given  $P$ -value can convey very different weights of evidence, in any *careful* intuitive sense, concerning a 'null hypothesis', on different occasions. I believe that this fact is leading, in statistics, to a new 'paradigm' (in the sense of Thomas Kuhn).

For other arguments in favor of  $W$ , and against other explications of weight of evidence, see Good (1988c, pp. 29 and 30).

<sup>8</sup> A private question from Suppes led to this and all earlier qualifications related to

indeterminate forms. He pointed out that if  $E^*$  is taken respectively as a 'tautology' or an impossibility (or, in the language of sets, 'the universe of possibilities' and the empty set), then (19) and (20), without qualification, would both lead neatly to a contradiction.

<sup>9</sup> I here use the 'compelling metaprinciple' of Good (1989).

<sup>10</sup> The pronounceable neologism *iff* is used here instead of the barbarism *iff*.

<sup>11</sup> It has been emphasized, for example, by Good (1975, p. 51) that it is much more difficult to judge the prior probability of a scientific theory than to judge the ratio of the prior probabilities of two rival scientific theories, although that too is by no means easy. That is why the concepts of generalized odds and generalized Bayes factors are important.

<sup>12</sup> The meaning of *state* is discussed by Good (1961/62, p. 51). To avoid a logical contradiction, I have to assume that  $U$  does not determine with physical certainty whether  $F$  occurs. In other words I am assuming that Nature is indeterministic, which is the fashionable opinion among physicists. For my views on this metaphysical matter see both indexes of Good (1983), and also Good (1988b).

## REFERENCES

- Fisher, R. A.: 1938, *Statistical Methods for Research Workers*, Oliver & Boyd, Edinburgh.
- Fisher, R. A.: 1956, *Statistical Methods and Scientific Inference*, Oliver & Boyd, Edinburgh.
- Good, I. J.: 1950, *Probability and the Weighing of Evidence*, Charles Griffin, London.
- Good, I. J.: 1961, 'Weight of Evidence, Causality, and False-Alarm Probabilities', in: Colin Cherry (Ed.), *Information Theory* (fourth London symposium, 1960), Butterworths, London, pp. 125-136.
- Good, I. J.: 1961/62, 'A Causal Calculus', *British Journal for the Philosophy of Science*, **11**, 305-318; **12**, 43-51; **13**, 88. Largely republished in Good (1983).
- Good, I. J.: 1965, 'Speculations Concerning the First Ultrainelligent Machine', *Advances in Computers*, **6**, Academic Press, New York, pp. 31-88.
- Good, I. J.: 1975, 'Explicativity, Corroboration, and the Relative Odds of Hypotheses', in a conference in the Dept. of Philosophy, Univ. of S. Carolina, 'Methodologies: Bayesian & Popperian', Nov. 1973, *Synthese*, **30**, 39-73. (Largely republished in Good, 1983.)
- Good, I. J.: 1983, *Good Thinking: The Foundations of Probability and its Applications*, Univ. of Minnesota Press.
- Good, I. J.: 1988a, 'Causal Tendency: A Review', in: Brian Skyrms and William L. Harper (Eds.), *Causation, Chance, and Credence*, proceedings of a conference at Irvine, California in 1985, Dordrecht, Kluwer, pp. 23-50.
- Good, I. J.: 1988b, 'Response to Wesley Salmon's Critique', in: William Harper and Brian Skyrms (Eds.), *Causation, Chance, and Credence*, Dordrecht, Reidel, pp. 73-78.
- Good, I. J.: 1988c, 'Changing Concepts of Chance', a review of *The Probabilistic Revolution*, Vol. 1, Ideas in History, L. Krüger, L.J. Daston, and M. Heidelberger

- (Eds.), Vol. 2, Ideas in the Sciences, L. Krüger, G. Gigerenzer, and Mary S. Morgan (Eds.), *Nature*, **332**, 1988, 405–406.
- Good, I. J.: 1989, 'Weight of Evidence and a Compelling Metaprinciple', *C319, J. Statist. Comput. & Simul.*, **31**, 121–123.
- Good, I. J.: 1991, 'The Bayes/Non-Bayes Compromise: A Brief Review. (The Longer Version)', Technical Report 91–16B, Statistics Department, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061–0439.
- Good, I. J.: 1992a, 'The Mathematics of Philosophy, A Brief Review of My Work', in: I. C. Jarvie (forthcoming). Essentially the same as Technical Report #92–3 (February 28), Statistics Department, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061.
- Good, I. J.: 1992b, 'The Bayes/Non-Bayes Compromise: A Brief Review', *Journal of the American Statistical Association* **87**, 597–602.
- Good, I. J.: 1992c, 'Introductory remarks for the article in *Biometrika*, **66** (1979), "A. M. Turing's Statistical Work in World War II", in: J. L. Britton (Ed.), *Collected Works of A. M. Turing*, volume entitled *Pure Mathematics: with a section on Turing's statistical work by I. J. Good*, North Holland, New York, pp. 211–223.
- Hart, H. L. A. and Honoré, A. M.: 1959, *Causation in the Law*, Clarendon Press, Oxford.
- Hilbert, D. and Ackermann, W.: 1950, *Principles of Mathematical Logic*, Chelsea Publishing Company, New York. Translated from the German of 1928 and 1938.
- Jarvie, I. C.: forthcoming, *The Enterprise of Critical Rationalism*, Kluwer, Dordrecht.
- Kempthorne, O. and Doerfler, T. E.: 1969, 'The Behavior of Some Significance Tests under Experimental Randomization', *Biometrika*, **56**, 231–248.
- Marx, Groucho: 1932, in the film *Horse Feathers*.
- Peirce, C. S.: 1878, 'The Probability of Induction', *Popular Science Monthly*; reprinted in: James R. Newman (Ed.), *The World of Mathematics 2*, Simon and Schuster, New York, 1956, pp. 1341–1354.
- Salmon, W.: 1984, *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton, New Jersey.
- Suppes, P.: 1970, *A Probabilistic Theory of Causality*, North-Holland Publishing Company, Amsterdam.
- Turing, A. M.: 1941, Personal communication.
- Wrinch, D. and Jeffreys, H.: 1921, 'On Certain Fundamental Principles of Scientific Discovery', *Philosophical Magazine*, Series 6, **42**, 369–390.

#### COMMENTS BY PATRICK SUPPES

Jack Good and I have been discussing causal notions for at least thirty years. My own work was very much influenced by his early publications (Good, 1961, 1962). Of course I can't claim to have read everything he has written even about causality and the weight of evidence because there is so much to look at in his continuous stream of papers. However, I did find the paper for this volume extremely interesting, and in my

view important, for its clear formulation of a notion of probabilistic sufficient cause and the companion concept of probabilistic necessary cause. (I am not happy with the new terminology of necessity and sufficientity he introduces, but I do think the concepts are important.)

As far as I know, in spite of the long tradition of emphasizing the difference between necessary and sufficient conditions, this standard distinction has not been introduced in an explicit way in the literature on probabilistic causality, at least not by me in my 1970 monograph or in any of the other literature with which I am familiar. Furthermore the definition that Jack gives makes important connections with the standard statistical concept of likelihood ratio, and thereby increases considerably the likelihood of statistical use of these concepts. As yet, it is fair to say, the literature on probabilistic causality has not had much impact on the actual practice of statistical inference about causal matters, although it has had an impact on the conceptual discussion of causality by statisticians. There is a chance that Good's new concepts will have a more direct influence.

Because Good introduces so much notation and has so many concepts at work, even though as he points out, the mathematical apparatus is elementary, I think it may be useful to abstract out the simplest cases and give the most direct definitions in terms of probability. I emphasize that the definitions I give are very much implicit in his paper but are not given at the beginning. For example, what I write for the definition of sufficient cause is almost an immediate consequence of his equation (13A). I use his notation of  $Q_{\text{suf}}(E : F | U)$  for  $F$  being a probabilistic sufficient cause for  $E$  with background  $U$ . The definition in terms of probability is then:

$$(1) \quad Q_{\text{suf}}(E : F | U) = \log \left[ \frac{P(\bar{E} | \bar{F}.U)}{P(\bar{E} | F.U)} \right]$$

The definition of probabilistic necessary cause is just obtained from (1) by negating every occurrence of  $E$  and  $F$ , which means that the negations in (1) will be replaced by double negations equal to the original expressions  $E$  and  $F$ . Thus we have

$$(2) \quad Q_{\text{nec}}(E : F | U) = \log \left[ \frac{P(E | F.U)}{P(E | \bar{F}.U)} \right]$$

Several nice examples are given by Good to illustrate these concepts intuitively. I mention just one. Let  $E$  be the event of being hit by

an automobile, and let  $F$  be the event of going for a walk. Then it is clear that the probabilistic necessary cause has a very high value but the sufficient cause does not. This seems correct, for if the sufficient cause were high, we would seldom go for a walk. On the other hand, we can recognize the high value of the necessary cause if the event of being hit by an automobile does occur. The notions  $Q_{\text{suf}}$  and  $Q_{\text{nec}}$  are additive, which is not surprising given the use of logarithms. It is also important to note that there is a corresponding 'multiplicative' set of concepts which do not use the logarithm, and which I express using a prime on  $Q$  rather than the more elaborate notation Good uses. The two equations are as follows:

$$(3) \quad Q'_{\text{suf}}(E : F | U) = \left[ \frac{P(\bar{E} | \bar{F}.U)}{P(\bar{E} | F.U)} \right]$$

$$(4) \quad Q'_{\text{nec}}(E : F | U) = \left[ \frac{P(E | F.U)}{P(E | \bar{F}.U)} \right]$$

The multiplicative necessary cause  $Q'_{\text{nec}}$  corresponds most closely to my own definition of prima facie cause, for as is obvious,  $Q'_{\text{nec}}(E : F.U) > 1$  iff  $P(E | F.U) > P(E | \bar{F}.U)$ . Using these same concepts,  $Q'_{\text{suf}}$  corresponds to the concept of  $\bar{F}$  being a prima facie cause of  $\bar{E}$ , rather than  $F$  being a prima facie cause of  $E$ .

But we can say more, for (3) is immediately equivalent to the following equation:

$$(5) \quad Q'_{\text{suf}}(E : F | U) = \left[ \frac{1 - P(E | \bar{F}.U)}{1 - P(E | F.U)} \right].$$

We can see at once from (5) that  $Q'_{\text{suf}}(E : F | U)$  has a value greater than 1 iff again  $F$  is a prima facie cause of  $E$ . The proof of this last assertion follows at once from the following inequalities:

$$\begin{aligned} \frac{1 - P(E | \bar{F}.U)}{1 - P(E | F.U)} &> 1 \\ \text{iff } 1 - P(E | \bar{F}.U) &> 1 - P(E | F.U) \\ \text{iff } P(E | F.U) &> P(E | \bar{F}.U) \\ \text{iff } P(E | F.U) &> P(E | U). \end{aligned}$$

Thus we have established the following proposition.

PROPOSITION 1.  $Q'_{\text{suf}}(E : F | U)$  and  $Q'_{\text{nec}}(E : F | U)$  are both greater than 1 iff  $F$  is a prima facie cause of  $E$ .

I note at once that this necessary and sufficient condition does not mean at all that  $Q'_{\text{suf}}$  and  $Q'_{\text{nec}}$  have the same value. It is obvious from Equations (4) and (5) that this is not the case. What it means is that they require, in order to really work in the correct direction, that they have as a necessary and sufficient condition that  $F$  is a prima facie cause of  $E$ . Their relative values depend very much upon the quantitative probabilities, a point that is important to Good's exposition. I should mention that the proposition I have just stated is formally close to the assertions stated by Good just before Equation (44), that is, that every pair of real values of  $Q_{\text{nec}}$  and  $Q_{\text{suf}}$  is possible, provided only that they are the same sign or both are zero. Note that his proposition deals not with  $Q'$  but with  $Q$ . I have just rearranged matters as shown here in order to bring out the direct connection with my own definition of prima facie cause.

What this also shows is that the distinction between necessary and sufficient cause that Good has introduced is an important quantitative distinction that I entirely missed in my original analysis and, as he says in the present paper, he did also. In my judgment whether we take the definitions either in the additive or multiplicative form, that is, with or without the logarithm, what is presented by Good is a new and important distinction in the theory of probabilistic causality. It should become a part of the standard literature.

#### REFERENCES

- Good, J.: 1961/62, 'A Causal Calculus', *British Journal for the Philosophy of Science*, 11, 305-316; 12, 43-51; 13, 88.  
 Suppes, P.: 1970, *A Probabilistic Theory of Causality*. Amsterdam: North Holland.