

A Problem in Cascaded Inference: Determining the Inferential Impact of Confirming and Conflicting Reports from Several Unreliable Sources¹

DAVID A. SCHUM²
Rice University

AND

CLINTON W. KELLY³
University of Michigan

The occurrence or nonoccurrence of an event D has inferential significance. We query n sources or sensors who make individual reports about the occurrence or nonoccurrence of D . Their reports are either mutually confirming or there is some pattern of conflict among them. In this paper we develop expressions termed adjusted likelihood ratios which prescribe the inferential or diagnostic impact of the joint confirming or conflicting reports from the n sources. These expressions combine information about the inferential impact of D (and its complement \bar{D}) with information about the reliability of each source. We only consider the case in which the reporting behavior of any subset of the sources is *not* itself an inferentially significant event. Appropriate independence and conditional independence assumptions are necessary. Our formulations of adjusted likelihood ratio are applicable to a variety of medical, legal, military, and other inferential tasks.

INTRODUCTION

In the most straightforward of probabilistic inference tasks one makes revisions of the relative likeliness of two or more world states or hypotheses on the basis of the established occurrence or nonoccurrence of events having inferential impact on these hypotheses. However, in a very large

¹ Special thanks are due Dr. Paul E. Pfeiffer and Dr. Robert M. Thrall, Mathematical Sciences Department, and Dr. Wesley M. DuCharme, Psychology Department, Rice University, for their kind assistance. Preparation of this paper was sponsored in part by the Advanced Research Projects Agency and the Office of Naval Research Grant Nos. NONR-N-00014-73C-0149 and NR-197-023.

² Requests for reprints should be sent to the senior author in care of Psychology Department, Rice University, Houston, Texas 77001.

³ Now at Decisions and Designs, Inc., McLean, Virginia 22101.

class of real-life probabilistic inference tasks the actual occurrence or non-occurrence of any inferential event may be conjectural. The reason is that the source or sources of information about this event may be less than perfectly reliable for any one of several reasons. Because there is often no single perfectly reliable source or sensor, two or more sources are asked to determine whether or not some inferential event has occurred. In X-ray interpretation, for example, several radiologists may be asked to determine whether or not a given X-ray image contains a certain indication. In legal proceedings, several witnesses may be asked to testify about the occurrence or nonoccurrence of some event. In intelligence analysis, more than one human source or more than one mechanical or electronic sensor may be used to determine whether or not some crucial event has occurred. The purpose of our paper is to develop formal statements which prescribe the inferential impact in the joint report, from two or more unreliable sources, about the occurrence or nonoccurrence of an inferential event.

Following is a summarization of the key concepts in our formalizations. Together, these concepts constitute a description of the basic inferential problems to which our formalizations are relevant. Because we are discussing only a particular aspect of cascaded inference, our notations can be kept relatively simple. We have chosen to keep our notational scheme consistent with the one adopted in an earlier paper on the relationships between the reliability of a single source and the inferential impact of the reports made by this source (Schum and Du Charme, 1971).

(1) Consider the single binomial event class $\{D, \bar{D}\}$. Suppose we are interested in determining the occurrence or nonoccurrence of event D for the inferential purpose of revising our estimate of the relative likeliness of two other events or hypotheses H_1 and H_2 . We assume that H_1 and H_2 are mutually exclusive but not necessarily exhaustive. A measure of the *inferential impact* of event D on H_1 and H_2 is provided by the likelihood ratio $L_D = P(D|H_1)/P(D|H_2) = p_1/p_2$. For event \bar{D} (nonoccurrence of event D), its *inferential impact* is given by the likelihood ratio $L_{\bar{D}} = P(\bar{D}|H_1)/P(\bar{D}|H_2) = (1 - p_1)/(1 - p_2)$.

(2) When a particular source or sensor, in making reports about events D and \bar{D} , is unreliable to any degree, a distinction must be made between the occurrence of event D and the *report* of the occurrence of event D . The necessity for such a distinction is discussed in a previous paper (Schum and Du Charme, 1971). A *report* of the occurrence of event D is to be symbolized by D^* and the *report* of the occurrence of event \bar{D} is to be symbolized by \bar{D}^* .

(3) We will consider a general case in which n sources (or sensors) are each asked to determine which one of two events (D or \bar{D}) has occurred. It is important to note that each source is being queried about the *same*

event pair $\{D, \bar{D}\}$. We let $N = \{1, 2, \dots, i, \dots, n\}$ be the set of integers which identify the sources of information about the occurrence of D or \bar{D} . Thus D^*_i is the report made by source i that event D occurred. Similarly, \bar{D}^*_i is the report from source i that event \bar{D} occurred. On occasion we shall refer to the class of reports $R = \{R_1, R_2, \dots, R_i, \dots, R_n\}$ where any R_i can be either D^*_i or \bar{D}^*_i . The subscript attached to any report thus indicates the source or sensor from which the report comes.

(4) We assume that no one of the n sources is perfectly reliable. That is, every source has, on previous occasions, reported D^* when event \bar{D} actually occurred and has reported \bar{D}^* when event D actually occurred. Specification of source reliability involves the following conditional probabilities which, for later convenience, are given labels which arise in signal detectability theory. The four conditional probabilities for any source i are:

- (a) $P(D^*_i|D) = h_i = \text{hit rate}$,
- (b) $P(\bar{D}^*_i|D) = 1 - h_i = m_i = \text{miss rate}$,
- (c) $P(\bar{D}^*_i|\bar{D}) = c_i = \text{correct rejection rate}$,
- (d) $P(D^*_i|\bar{D}) = 1 - c_i = f_i = \text{false alarm rate}$.

Our assumption about imperfect source reliability means that $h_i \neq 1.0$ and that $c_i \neq 1.0$, for every source i . For reasons which will become apparent when we discuss all-important independence considerations, the best indications of "reliability" for source i are provided by the ratios h_i/f_i and c_i/m_i . There is a final consideration about source reliability. In some cases a given source might express the extent of his own uncertainty about the occurrence of events D and \bar{D} . For example, he might say that he is 60% sure that D occurred and 40% sure that \bar{D} occurred. The formalizations we develop do not consider this type of reporting (see Gettys and Willke, 1969). We are concerned instead with instances in which every source unequivocally reports D^* or \bar{D}^* . The uncertainty about the actual occurrence of events D or \bar{D} must be estimated or expressed on the basis of a source's previous reporting accuracy. The four conditional probabilities listed above are formally required for this purpose.

(5) When a source of information about D or \bar{D} is less than completely reliable, an inference about the relative likeliness of H_1 and H_2 should be based upon reports D^* or \bar{D}^* having adjusted likelihood ratios $\Lambda_{D^*} = P(D^*|H_1)/P(D^*|H_2)$ and $\Lambda_{\bar{D}^*} = P(\bar{D}^*|H_1)/P(\bar{D}^*|H_2)$. In their expanded form these "adjusted likelihood ratios" combine information about source reliability (expressed by h_i/f_i and c_i/m_i) with information about the impact of the event being reported (expressed by the values p_1 and p_2).

(6) A crucial factor in our formalizations is the particular pattern of reports made by the n sources about events D and \bar{D} . First, suppose that

all n sources make the same report (i.e., they all make reports D^* or they all make reports \bar{D}^*). We shall call this the *confirming* case since all sources are consistent in making the same report. In the second case, suppose that $r \geq 1$ of the sources make reports D^* but the remaining $(n - r)$ sources make reports \bar{D}^* . This case will be called the *conflicting* case. In the present paper we develop formal statements which prescribe the inferential or diagnostic impact in the *joint report*, from two or more sources, of events in the class $\{D, \bar{D}\}$. We consider both the confirming and conflicting cases. Each of the statements we develop is in the form of an adjusted likelihood ratio (Λ). A value of Λ will combine information about the reliability of each source with information about the inferential impact of events in the class $\{D, \bar{D}\}$.

MAJOR INDEPENDENCE ISSUES

We begin by considering ways of characterizing the individual and the joint behavior of sources or sensors making reports D^* or \bar{D}^* . The concepts of *independence* and *conditional independence* are crucial in these considerations. We recall that events A and B are said to be (stochastically) independent iff (if and only if) $P(AB) = P(A)P(B)$. If this relationship holds, then $P(A|B) = P(A)$ and $P(B|A) = P(B)$. In this case it is frequently said that A and B are *totally independent*. Events A and B are said to be *conditionally independent*, given another event C , iff $P(AB|C) = P(A|C)P(B|C)$. It can easily be demonstrated that conditional independence of A and B (given C) does not imply total independence of A and B nor does total independence imply conditional independence. Our present problem of determining the inferential impact of a joint report from unreliable sources will be a prime example showing the absolute necessity of making careful and separate independence assumptions about various aspects of source behavior. Indeed, lack of precision in stating the independence conditions we wish to assume will lead us to conclusions about adjusted likelihood ratio whose consequences are not at all congenial to intuition in a variety of conditions. There is a large class of independence issues of concern in cascaded inference (Schum and Du Charme, 1971). We now list two independence issues of importance in describing the individual and joint behavior of our reporting sources. Independence issues crucial in other stages of cascaded inference are discussed in a later section.

1. *Report and Event Nonindependence*

Our first formal requirement is that the individual reports from each source actually convey information about the occurrence or nonoccurrence of event D . This requirement can be expressed as an assumption about the total *nonindependence* of reports (D^*) and events (D).

Assumption 1. Consider the class $\{D^*_i, D\}$ where i is any source in $N = \{1, 2, \dots, n\}$. We assume that the class $\{D^*_i, D\}$ is *not* an independent class for any $i = 1, 2, \dots, n$. We note that this assumption implies that the class $\{\bar{D}^*_i, \bar{D}\}$ is also *not* an independent class for any $i = 1, 2, \dots, n$.

Statistical independence can be interpreted as a failure of conditioning of one event by another (Pfeiffer and Schum, 1973). Assumption 1 simply states that reports from any single source *are* conditional upon the events being reported; that is, $P(D^*_i|D) \neq P(D^*)$ and $P(\bar{D}^*_i|\bar{D}) \neq P(\bar{D}^*)$. An easily verified implication of Assumption 1 is that $P(D^*_i|D) = h_i \neq f_i = P(D^*_i|\bar{D})$, and that $P(\bar{D}^*_i|\bar{D}) = c_i \neq m_i = P(\bar{D}^*_i|D)$. It has been shown that, with the exception of certain unusual cases discussed elsewhere, the condition $h_i = f_i$ (implying $c_i = m_i$) for any source i causes a total destruction of the inferential impact of reports from this source (Schum and Du Charme, 1971). It can be argued, therefore, that the condition $h_i/f_i = 1.0$ might best be taken as the condition of zero reliability for source i .

In our formalizations of adjusted likelihood ratio (Λ) in both cases we will assume that $h_i > f_i$ (which implies that $c_i > m_i$). The reason for this assumption is that $h_i > f_i$ means that the report D^* will favor inferentially the same hypothesis that event D favors (the same can be said about \bar{D}^* and \bar{D}). If $h_i < f_i$ (or $c_i < m_i$) a condition called an *impact reversal* occurs. In an impact reversal a source report favors one hypothesis while the *event* being reported favors another hypothesis. We simply wish to exclude these possible impact reversals in our developments. It does seem likely that few real-life inference systems would employ sources whose hit rates might be smaller than their false-alarm rates. Impact reversals are possible but seem pathological or unlikely in most inferential activity.

2. Source Behavior, by Itself, Has No Inferential Significance

Our second assumption is more subtle than the first and it requires careful statement. We wish to consider situations in which all of the inferential impact upon H_1 and H_2 resides in the events $\{D, \bar{D}\}$ being reported upon and *not in the individual or joint behavior of the sources or sensors making reports about D and \bar{D}* . The expressions we develop for Λ will tell us how much of the total impact in events D and \bar{D} [expressed by $L_D = p_1/p_2$ and $L_{\bar{D}} = (1 - p_1)/(1 - p_2)$] we should incorporate in our inferences about the relative likeliness of H_1 and H_2 . Exactly how much of this total impact we can incorporate will depend upon the reliability of the sources making reports D^* and \bar{D}^* . When source behavior, by itself, has no inferential significance we can expect expressions for Λ to be appro-

privately bounded by L_D and $L_{\bar{D}}$ in both the confirming and conflicting cases.

There are two varieties of source behavior we wish to rule out. A single appropriate conditional independence assumption will allow us to make these exclusions. First, we wish to specify that the reporting behavior of any subset of sources (including single-member subsets) is not contingent upon any hypothesis about which inferences are being made. For example, we exclude possible instances in which a radiologist's hit or false-alarm rate might depend upon what illness (H_i) a patient has. Second, we will require the source reports to be *mutually* independent given event D and given event \bar{D} . As an example, suppose five radiologists read and interpret X-ray images for us. We will require that the reporting behavior of any subset (single-member subsets included) of the radiologists is not contingent upon the reporting behavior of any other subset, both when event D is true and when event \bar{D} is true. The reason for requiring mutual source independence *given D and given \bar{D}* is that, by Assumption 1, the reports from each *individual* source *are* to be conditional upon D and \bar{D} .

It may appear that we must make two separate assumptions. However, the following statement incorporates both exclusions we wish to make.

Assumption 2. Consider the class $E = \{R_1, R_2, \dots, R_n, H_k\}$ where any $R_i = D^*_i$ or \bar{D}^*_i , and $k = 1, 2$. We assume that class E is *conditionally independent, given D and given \bar{D}* , for H_1 and H_2 . For any subclass of the R_i and for $k = 1, 2$, this assures:

$$P\left(\bigcap_i R_i | D \cap H_k\right) = P\left(\bigcap_i R_i | D\right) = \prod_i P(R_i | D), \text{ and}$$

$$P\left(\bigcap_i R_i | \bar{D} \cap H_k\right) = P\left(\bigcap_i R_i | \bar{D}\right) = \prod_i P(R_i | \bar{D}).$$

Reading both of the above expressions from left to right we first observe the provision that H_k ($k = 1, 2$) fails to condition either $P(\bigcap_i R_i | D)$ or $P(\bigcap_i R_i | \bar{D})$ for any subset of the R_i . As an illustration of the consequences of not making this assumption, suppose that $P(\bigcap_i R_i | D \cap H_k) \neq P(\bigcap_i R_i | D)$ for some subset of the R_i . When this is true, it can easily be deduced that $P(H_k | \bigcap_i R_i \cap D) \neq P(H_k | D)$. In words, this inequality means that there is inferential impact on H_k of the joint report $(\bigcap_i R_i)$ *in excess* of the impact in the event D being reported. As mentioned above, in our formalizations of adjusted likelihood ratio we wish to exclude instances in which the reporting behavior of one or more of the sources is actually an event itself having inferential significance regarding H_1 and H_2 . Reporting behavior as an inferential event is a problem which appears to be of some

importance in certain inferential systems and it is briefly discussed for single sources in Schum and Du Charme (1971).

The second provision is expressed by the product rule for conditional probabilities in each of the above equality expressions. These product rules embody the assumption of source mutual independence, given D and given \bar{D} . To illustrate consequences of the failure of this assumption, consider three sources $i = 1, 2, 3$, and suppose $P(\bigcap_{i=1}^3 R_i|D) \neq \prod_{i=1}^3 P(R_i|D)$.

One particular instance in which this product rule may fail to hold is the case in which $P(R_1 \cap R_2|D \cap R_3) \neq P(R_1 \cap R_2|D)$. This latter inequality implies that $P(R_3|D \cap R_1 \cap R_2) \neq P(R_3|D)$. Suppose $R_3 = D^*_3$, then $P(D^*_3|D \cap R_1 \cap R_2) \neq P(D^*_3|D) = h_3$. In other words, the hit rate for Source 3 is influenced by reports R_1 and R_2 from Sources 1 and 2 in this case. It is this sort of mutual influence among the sources that we wish to exclude.

Finally, we note that Assumption 2 does *not* imply the *complete* or *unconditional* independence of the R_i . In other words, we cannot deduce from Assumption 2 that $P(\bigcap_i R_i) = \prod_i P(R_i)$, for any subset of the R_i . This is not crucial in our development of Λ because, under Assumption 1, we suppose that every R_i is conditioned by D and \bar{D} .

ADJUSTED LIKELIHOOD RATIO: CONFIRMING CASE

We first consider the case in which there is complete reporting agreement among the n sources. All n sources report the same condition, i.e., every $R_i = D^*_i$ or every $R_i = \bar{D}^*_i$. We recall our initial provision that, for any source i , $h_i \neq 1.0$ and $c_i \neq 1.0$. Our task is to find expressions which prescribe the inferential impact, on the mutually exclusive hypotheses H_1 and H_2 , of the joint confirming reports from the n sources. We begin with the following definition.

Definition 1. Suppose $R_i = D^*_i$, for every source i . The inferential impact on H_1 and H_2 of these n confirming reports is prescribed by the *adjusted likelihood ratio*

$$\Lambda[n(D^*)] = P \left[\bigcap_{i=1}^n D^*_i | H_1 \right] / P \left[\bigcap_{i=1}^n D^*_i | H_2 \right].$$

If $R_i = \bar{D}^*_i$, for every source i , the inferential impact on H_1 and H_2 of these confirming reports is prescribed by the *adjusted likelihood ratio*

$$\Lambda[n(\bar{D}^*)] = P \left[\bigcap_{i=1}^n \bar{D}^*_i | H_1 \right] / P \left[\bigcap_{i=1}^n \bar{D}^*_i | H_2 \right].$$

To illustrate the development of adjusted likelihood ratio for confirming reports we restrict our attention to $\Lambda[n(D^*)]$. The development of $\Lambda[n(\bar{D}^*)]$ is similar in all respects. For simplification we let $F = \bigcap_{i=1}^n D^*_i$. For any k

we can write: $P[\bigcap_{i=1}^n D^*_i|H_k] = P(F|H_k) = P(F \cap H_k)/P(H_k)$. Then,

$$\begin{aligned} P(F \cap H_k) &= P(F \cap D \cap H_k) + P(F \cap \bar{D} \cap H_k) \\ &= P(F|D \cap H_k)P(D|H_k)P(H_k) + P(F|\bar{D} \cap H_k)P(\bar{D}|H_k)P(H_k). \end{aligned}$$

Then, by definition,

$$P(F|H_k) = P(F|D \cap H_k)P(D|H_k) + P(F|\bar{D} \cap H_k)P(\bar{D}|H_k).$$

Under Assumption 2 we can write: $P(F|H_k) = P(F|D)P(D|H_k) + P(F|\bar{D})P(\bar{D}|H_k)$

$$\begin{aligned} P(\bar{D}|H_k). \text{ Also under Assumption 2, } P(F|D) &= \prod_{i=1}^n P(D^*_i|D) \text{ and } P(F|\bar{D}) \\ &= \prod_{i=1}^n P(D^*_i|\bar{D}). \text{ Thus,} \end{aligned}$$

$$P\left(\bigcap_{i=1}^n D^*_i|H_k\right) = \prod_{i=1}^n P(D^*_i|D)P(D|H_k) + \prod_{i=1}^n P(D^*_i|\bar{D})P(\bar{D}|H_k).$$

This development holds for any k , so we can write:

$$\Lambda[n(D^*)] = \frac{\left[\prod_{i=1}^n P(D^*_i|D) \right] P(D|H_1) + \left[\prod_{i=1}^n P(D^*_i|\bar{D}) \right] P(\bar{D}|H_1)}{\left[\prod_{i=1}^n P(D^*_i|D) \right] P(D|H_2) + \left[\prod_{i=1}^n P(D^*_i|\bar{D}) \right] P(\bar{D}|H_2)}. \tag{1}$$

Various alternative forms of Eq. (1) will prove useful. We recall that $P(D^*_i|D) = h_i$ (the ‘‘hit’’ rate for source i) and that $P(D^*_i|\bar{D}) = f_i$ (the ‘‘false alarm’’ rate for source i). We let $A_i = h_i/f_i$ be the hit rate relative to the false-alarm rate for source i . Also, let $p_1 = P(D|H_1)$ and $p_2 = P(D|H_2)$, $P(\bar{D}|H_1) = 1 - p_1$, and $P(\bar{D}|H_2) = 1 - p_2$. Then we can express $\Lambda[n(D^*)]$ alternatively as:

$$\Lambda[n(D^*)] = \frac{\left[\prod_{i=1}^n A_i \right] p_1 + (1 - p_1)}{\left[\prod_{i=1}^n A_i \right] p_2 + (1 - p_2)} \tag{1a}$$

$$= \frac{p_1 \left[\prod_{i=1}^n A_i - 1 \right] + 1}{p_2 \left[\prod_{i=1}^n A_i - 1 \right] + 1} \quad (1b)$$

$$= \frac{p_1 + \alpha}{p_2 + \alpha}, \quad \text{where } \alpha = \left[\prod_{i=1}^n A_i - 1 \right]^{-1}, \quad (1c)$$

provided that $\prod_{i=1}^n A_i \neq 1$. In all discussions of $\Lambda[n(D^*)]$ we assume that $h_i > f_i$ for any source i (i.e., $A_i > 1$, all i). In Eq. (1c) this means that $0 < \alpha < \infty$.

Using the same development and assumptions we can express $\Lambda[n(\bar{D}^*)]$ as:

$$\Lambda[n(\bar{D}^*)] = \frac{\left[\prod_{i=1}^n P(\bar{D}^*_i|\bar{D}) \right] P(\bar{D}|H_1) + \left[\prod_{i=1}^n P(\bar{D}^*_i|D) \right] P(D|H_1)}{\left[\prod_{i=1}^n P(\bar{D}^*_i|\bar{D}) \right] P(\bar{D}|H_2) + \left[\prod_{i=1}^n P(\bar{D}^*_i|D) \right] P(D|H_2)}. \quad (2)$$

We recall that $P(\bar{D}^*_i|\bar{D}) = c_i$ (the "correct" rejection rate for source i) and that $P(\bar{D}^*_i|D) = m_i$ (the "miss" rate for source i). We let $B_i = c_i/m_i$ indicate the correct rejection rate relative to the miss rate for source i . Then, alternatively,

$$\Lambda[n(\bar{D}^*)] = \frac{\left[\prod_{i=1}^n B_i \right] (1 - p_1) + p_1}{\left[\prod_{i=1}^n B_i \right] (1 - p_2) + p_2} \quad (2a)$$

$$= \frac{p_1 + \beta}{p_2 + \beta}, \quad \text{where } \beta = \prod_{i=1}^n B_i / 1 - \prod_{i=1}^n B_i \quad (2b)$$

provided that $\prod_{i=1}^n B_i \neq 1$. The previous requirement $h_i > f_i$ implies that $c_i > m_i$ (i.e., $B_i > 1$, for every i). This means that $-\infty < \beta < -1$ in Eq. (2b). We recall that the reason for requiring $A_i > 1$ (all i) in $\Lambda[n(D^*)]$ and $B_i > 1$ (all i) in $\Lambda[n(\bar{D}^*)]$ is to rule out "pathological" instances in which an event points inferentially to one hypothesis but the report of this event points inferentially to the other hypothesis. Any source whose

hit rate is less than its false alarm rate can cause “impact reversals” of the sort mentioned previously.

Discussion of the Confirming Case

In the following comments we describe various properties of adjusted likelihood ratio in the confirming case. The first comment should assist the reader in determining why special care in making assumptions about independence and conditional independence was necessary.

1. *Limiting Values of $\Lambda[n(D^*)]$ and $\Lambda[n(\bar{D}^*)]$*

Using Eq. (1c), we observe that, if $A_i > 1$ for all i , then $\Lambda[n(D^*)]$ is bounded by the values 1 and p_1/p_2 . As $\prod_{i=1}^n A_i$ is made larger, α goes to zero and $\Lambda[n(D^*)]$ approaches p_1/p_2 . As $\prod_{i=1}^n A_i$ approaches 1, α increases

without limit and $\Lambda[n(D^*)]$ approaches 1. Thus, if $p_1 > p_2$, then $1 < \Lambda[n(D^*)] < p_1/p_2$. If $p_1 < p_2$, then $p_1/p_2 < \Lambda[n(D^*)] < 1$. Similar arguments show that $\Lambda[n(\bar{D}^*)]$ is bounded by the values 1 and $1 - p_1/1 - p_2$ when $B_i > 1$, for every i . When $p_1 > p_2$, then $(1 - p_1)/(1 - p_2) < \Lambda[n(\bar{D}^*)] < 1$. When $p_1 < p_2$, then $1 < \Lambda[n(\bar{D}^*)] < (1 - p_1)/(1 - p_2)$.

We now ask why we should expect these particular bounds on adjusted likelihood ratio in the confirming case given our assumptions. Suppose the n sources (each one not completely reliable) agree and make reports D^* . These n confirming reports all refer to the occurrence of the same event D . If D actually occurred, its total inferential impact on H_1 and H_2 is indicated by the ratio p_1/p_2 . Now, our Assumption 2, that reporting behavior given D and \bar{D} is not contingent upon any H_k , rules out instances in which adjusted likelihood ratio could, legitimately, provide impact over and above the impact of event D . Recall the example showing that, when any R_i is conditioned by any H_k , the reporting behavior of the sources has inferential significance itself. We are requiring also that the joint confirming report should provide no inferential impact in excess of the impact of event D . We wish our models to reflect the fact that all n reports refer to the same event and that no report or subset of reports contain inferential information not already contained in event D . The model for $\Lambda(n(D^*))$ shown above is consistent in this respect. $\Lambda[n(D^*)]$ prescribes how much of the total impact of event D (on H_1 and H_2) we should apply in revising our opinions about the relative likeliness of H_1 and H_2 on the basis of the joint confirming report. How much of this total impact we can apply depends upon the reliability of the sources.

Now, with only slight modification, we could have stated Assumption 2 in a manner which leads to entirely different consequences in both the confirming and conflicting cases. Suppose we had simply required that the report class $R = \{R_1, R_2, \dots, R_n\}$ be independent conditional upon H_k , for $k = 1, 2$. Under this assumption we can write, for the case in which all sources report D^* ,

$$\Lambda[n(D^*)] = \frac{P\left[\bigcap_{i=1}^n D^*_i | H_1\right]}{P\left[\bigcap_{i=1}^n D^*_i | H_2\right]} = \prod_{i=1}^n \frac{[P(D^*_i | H_1)]}{[P(D^*_i | H_2)]} = \prod_{i=1}^n \Lambda_{D^*_i},$$

where $\Lambda_{D^*_i}$ is the adjusted likelihood ratio for the report D^* from the single source i .

It can easily be shown that $\prod_{i=1}^n \Lambda_{D^*_i}$ is *not* bounded by values prescribing the impact of *event* D , even when $h_i < 1$ and $c_i < 1$ for every source i . The alternative assumption stated in the previous paragraph does not imply the requirement in our original Assumption 2 that the R_i be *mutually independent given D and given \bar{D}* , under h_i and H_2 . Thus, our Assumption 2 also rules out instances in which there may be inferential impact in the joint reporting behavior of the sources which is not due to the influence of H_1 or H_2 . The assumption leading to the condition $\Lambda[n(D^*)] = \prod_{i=1}^n \Lambda_{D^*_i}$ fails to rule out this possible additional impact and so we expect that adjusted likelihood ratio would not be bounded by the total impact of events $\{D, \bar{D}\}$.

2. Contingency of $\Lambda[n(D^*)]$ and $\Lambda[n(\bar{D}^*)]$ upon Specific Values of p_1 and p_2

Equations (1c) and (2b) make clear the functional dependence of $\Lambda[n(D^*)]$ and $\Lambda[n(\bar{D}^*)]$ upon *specific* values of p_1 and p_2 . This means, for example, that in prescribing the impact of *event* D , we must have $P(D|H_1) = p_1$ and $P(D|H_2) = p_2$. Merely prescribing the likelihood ratio $L_D = P(D|H_1)/P(D|H_2) = p_1/p_2$ for event D will not suffice. Although prescribing the ratio L_D may be an easier behavioral task in many instances, L_D suppresses information about how probable the event is under both H_1 and H_2 . It happens that the extent to which source unreliability degrades inferential impact depends upon how probable an event is under both H_1 and H_2 . Figure 1 shows $\Lambda[n(D^*)]$ as a function of α for two events D_1 and D_2 . For event D_1 , $p_1 = 0.9$ and $p_2 = 0.09$. For event D_2 , $p_1 = 0.2$ and $p_2 = 0.02$. Notice that, although $L_{D_1} = L_{D_2} = 10$, event D_2 is more im-

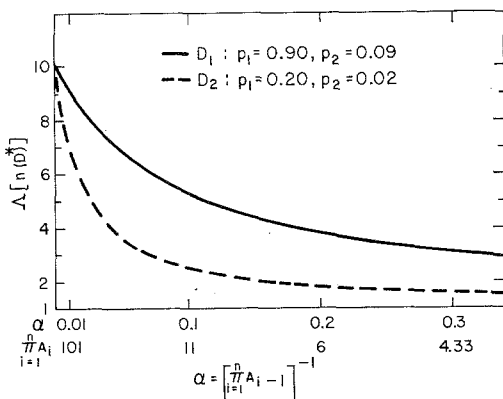


FIG. 1. $\Delta[n(D^*)]$ as a function of α in two special cases.

probable than event D_1 under H_1 and H_2 . For every $\alpha > 0$, the inferential impact in the joint report of D_2 is less than the impact in the joint report of D_1 . In general, the more improbable an event is under H_1 and H_2 the more seriously degraded is the inferential impact in reports of this event by sources having some fixed level of unreliability.

3. Adding Additional Reporting Sources

We can consider the n sources to be a reporting or detection "system." Under our particular set of assumptions $\prod_{i=1}^n A_i$ represents the combined

hit rate relative to false alarm rate of this entire system. Similarly, $\prod_{i=1}^n B_i$

represents the combined correct rejection rate relative to the miss rate for the system. Recall our assumption that $A_i > 1$ (which implies $B_i > 1$) for every source i . Using a representation like Figure 1 one can easily determine the inferential effect of adding sources with known A_i if the reports from these additional sources are confirming. For example, suppose

we have $n = 3$ sources with $\prod_{i=1}^3 A_i = 11$, i.e., $\alpha = 0.1$. All three sources report the occurrence of event D whose p_1 and p_2 values are those for event D_1 in Figure 1. We find that $\Delta[3(D^*)] = 5.26$. We decide to query

an additional source with $A_4 = 4$. Now, $\prod_{i=1}^4 A_i = 44$ and $\alpha = 0.023$. If the additional source agrees with the other three sources, $\Delta[4(D^*)] = 8.17$.

ADJUSTED LIKELIHOOD RATIO: CONFLICTING CASE

Again, let $N = \{1, 2, \dots, n\}$ be the set of integers which identify the n sources of information about the occurrence or nonoccurrence of D . Let the class $\{I, J\}$ represent *any* partition of N such that set I contains exactly r members (where $1 \leq r < n$) and set J contains exactly $(n - r)$ members. In the conflicting case we are concerned with the joint occurrence of reports from two classes: the class $\{D^*_i; i \in I\}$ and the class $\{\bar{D}^*_j; j \in J\}$. Members of index set I represent the r sources which report D^* and members of index set J represent the $(n - r)$ sources which report \bar{D}^* . For any source $i \in I$ we assume $h_i/f_i = A_i > 1$, and for any source $j \in J$ we assume $c_j/m_j = B_j > 1$.

Definition 2. Suppose I represents the r sources making reports D^* and J represents the $n - r$ sources making reports \bar{D}^* , where $1 \leq r < n$. The inferential impact on H_1 and H_2 of the joint *conflicting* reports from these n sources is prescribed by the *adjusted likelihood ratio*:

$$\Lambda(I, J) = \frac{P \left[\left(\bigcap_{i \in I} D^*_i \right) \cap \left(\bigcap_{j \in J} \bar{D}^*_j \right) \mid H_1 \right]}{P \left[\left(\bigcap_{i \in I} D^*_i \right) \cap \left(\bigcap_{j \in J} \bar{D}^*_j \right) \mid H_2 \right]}$$

We continue to make Assumptions 1 and 2 as stated at the outset. We will simply list alternative forms of $\Lambda(I, J)$ since its development proceeds in exactly the same manner as the development of $\Lambda[n(D^*)]$. First, we have:

$$\Lambda(I, J) = \frac{\left[\prod_{i \in I} \frac{P(D^*_i \mid D)}{P(D^*_i \mid \bar{D})} \right] P(D \mid H_1) + \left[\prod_{j \in J} \frac{P(\bar{D}^*_j \mid \bar{D})}{P(\bar{D}^*_j \mid D)} \right] P(\bar{D} \mid H_1)}{\left[\prod_{i \in I} \frac{P(D^*_i \mid D)}{P(D^*_i \mid \bar{D})} \right] P(D \mid H_2) + \left[\prod_{j \in J} \frac{P(\bar{D}^*_j \mid \bar{D})}{P(\bar{D}^*_j \mid D)} \right] P(\bar{D} \mid H_2)} \tag{3}$$

With notational conventions already established we can write:

$$\Lambda(I, J) = \frac{\left[\prod_{i \in I} A_i \right] p_1 + \left[\prod_{j \in J} B_j \right] (1 - p_1)}{\left[\prod_{i \in I} A_i \right] p_2 + \left[\prod_{j \in J} B_j \right] (1 - p_2)} \tag{3a}$$

In the interests of further simplification we let $A_I = \prod_{i \in I} A_i$ and $B_J =$

$\prod_{j \in J} B_j$. That is, A_I is the combined hit rate relative to false alarm rate for all r sources reporting D^* and B_J is the combined correct rejection rate relative to miss rate for all $(n - r)$ sources reporting \bar{D}^* . Thus,

$$\Lambda(I, J) = \frac{A_I p_1 + B_J(1 - p_1)}{A_I p_2 + B_J(1 - p_2)} \quad (3b)$$

$$= \frac{p_1[(A_I/B_J - 1) + 1]}{p_2[(A_I/B_J - 1) + 1]} \quad (3c)$$

$$= (p_1 + \gamma)/(p_2 + \gamma), \quad (3d)$$

where

$$\gamma = [(A_I/B_J) - 1]^{-1},$$

provided that $A_I \neq B_J$.

Discussion of the Conflicting Case

1. *Inferential Direction and Strength of $\Lambda(I, J)$.*

Using Eq. (3d) we can establish boundary conditions for $\Lambda(I, J)$. If we fix B_J and let A_I increase without limit, γ approaches zero and $\Lambda(I, J)$ approaches p_1/p_2 . Fixing A_I and letting B_J increase without limit makes γ approach the value -1 and $\Lambda(I, J)$ approach the ratio $(1 - p_1)/(1 - p_2)$. When $p_1 > p_2$, then $(1 - p_1)/(1 - p_2) < \Lambda(I, J) < p_1/p_2$; when $p_1 < p_2$, then $p_1/p_2 < \Lambda(I, J) < (1 - p_1)/(1 - p_2)$. Thus, subject to the assumptions made at the outset, the inferential impact of the joint conflicting report is bounded by the impact of event D (prescribed by p_1/p_2) and the impact of \bar{D} (prescribed by $(1 - p_1)/(1 - p_2)$).

We must now consider the problem of finding the *inferential direction* of a joint conflicting report. By specifying the inferential "direction" of any report we simply mean specifying the hypothesis (H_1 or H_2) under which the report is most likely. We say a report "points toward" or "favors" the hypothesis under which the report is most likely. The inferential direction of a joint conflicting report depends upon A_I and B_J as well as upon p_1 and p_2 . We first note, using Eq. (3c), that when $A_I = B_J$, then $\Lambda(I, J) = 1$ regardless of the value of p_1 and p_2 . Thus, a joint conflicting report has no inferential impact when $A_I = B_J$. Now, when $A_I > B_J$, the joint conflicting report will favor whichever hypothesis event D favors. (Recall that the r sources, with combined hit rate relative to false alarm rate A_I , all reported the occurrence of event D .) When $A_I < B_J$, then the joint conflicting report will favor whichever hypothesis event \bar{D} favors. (The $n - r$ sources, with combined correct rejection rate to miss rate B_J , all reported the occurrence of event \bar{D} .) From Eq. (3c) we can see that $A_I > B_J$ makes $[(A_I/B_J) - 1] > 0$ so that $\Lambda(I, J) > 1$ when $p_1/p_2 > 1$ and $\Lambda(I, J) < 1$ when $p_1/p_2 < 1$. Thus, when $A_I > B_J$, the inferential direction of $\Lambda(I, J)$ is the same as the inferential direction of event D . Similarly, when $B_J > A_I$ in Eq. (3c), then $[(A_I/B_J) - 1] < 0$.

This means that $\Lambda(I, J) > 1$ when $p_2 > p_1$ (i.e., $(1 - p_1)/(1 - p_2) > 1$) and $\Lambda(I, J) < 1$ when $p_2 < p_1$ (i.e., $(1 - p_1)/(1 - p_2) < 1$). Thus, when $B_J > A_I$, $\Lambda(I, J)$ points inferentially to the same hypothesis as does event \bar{D} .

2. "Majority Rule" by Itself is Inappropriate as an Indication of Inferential Direction in the Conflicting Case

Equation (3d) makes it clear that the relative sizes of $A_I = \prod_{i \in I} A_i$ and $B_J = \prod_{j \in J} B_j$ determine the inferential direction of joint conflicting reports rather than the numbers of members r in I and $n - r$ in J . One can easily show that $r > n - r$ does not imply $A_I > B_J$, and, that $r < n - r$ does not imply $A_I < B_J$. The result is that one should "side" inferentially with the majority report iff the joint reliability of the majority sources exceeds the joint reliability of the sources in the minority. Thus, source reliability is more important than the extent of source agreement in our assessment of the inferential strength and direction of conflicting reports. This means that the practice of using an odd number of sources to resolve potential reporting conflicts ignores what is crucial; namely, source reliability. It also means that some thought should be given, in civil judicial proceedings, about whether or not the concept of "preponderance of evidence" is interpreted correctly. The number of witnesses testifying is not as crucial as their reliabilities.

3. An Example of Impact Determination in the Conflicting Case

Figure 2 shows $\log_{10} \Lambda[I, J]$ as a function of γ for the special case in which $P(D|H_1) = p_1 = 0.60$ and $P(D|H_2) = p_2 = 0.15$. $\log_{10} \Lambda[I, J]$ is plotted in an effort to remove some of the natural asymmetry of Λ values about 1.0, the point of zero inferential impact. Notice that γ is undefined in

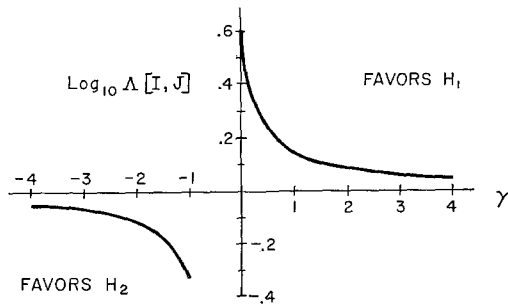


FIG. 2. $\text{Log}_{10} \Lambda[I, J]$ as a function of α in the special case $p_1 = 0.60$ and $p_2 = 0.15$.

$(-1, 0)$. This arises because both A_I and B_J are positive values. In this special case event D favors H_1 and event \bar{D} favors H_2 . A joint conflicting report $\log_{10} \Lambda[I, J] > 0$ favors H_1 . Such values occur when $\gamma \geq 0$ (i.e., when $A_I > B_J$). A joint conflicting report with $\log_{10} \Lambda[I, J] < 0$ favors H_2 . These values occur when $\gamma \leq -1$ (i.e., when $A_I < B_J$). A representation such as Figure 2 can be used to find the change in inferential strength and direction of a joint conflicting report when additional sources, each with known values of A_i or B_j , are added. We note that in the conflicting case, as well as in the confirming case, that Λ depends upon p_1 and p_2 rather than upon the ratio of these values.

OTHER CONDITIONAL INDEPENDENCE ISSUES

Independence issues arise, other than those previously mentioned, when we attempt to evaluate a current joint report in the light of previous evidence. In most inference situations one's evidence or knowledge, prior to the acquisition of a current report or observation, consists of a melange of verified and unverified reports. Our task is to determine the extent to which any of this prior knowledge should further condition the information we are now using to assess the impact of our current joint report. The models for adjusted likelihood ratio in both the confirming and conflicting cases make it clear that we need two classes of information in order to assess the inferential impact of a joint report. The first kind of information concerns the inferential impact of the *event* being reported. This impact is specified by the values p_1 and p_2 . The second type of information concerns the reliability of the sources or observers making reports about the occurrence or nonoccurrence of the current event of concern. The reliability of a source is indicated by its hit and false-alarm rates. Our specific concern is with cases in which values indicating event impact and source reliability have some probabilistic linkage with prior evidence. The following comments are intended merely to alert the reader to the conditional independence issues involved.

Let T represent the class of all relevant events E_t known to have occurred prior to a current joint report. Let U represent the class of relevant *unverified* reports R_u from sources with less than perfect reliability. By *relevant* we mean that these events and reports have impact on inferences involving two disjoint hypotheses H_1 and H_2 . We let S represent our entire set of prior evidence, where

$$S = \left[\bigcap_{t \in T} E_t \right] \cap \left[\bigcap_{u \in U} R_u \right].$$

Suppose the class $\{D_y, \bar{D}_y\}$ includes a *current* event and its complement, and the class $\{D^*_{u_i}, \bar{D}^*_{u_i}\}$ consists of reports of either of these two current

events made by source i (whose reliability is less than perfect). Further, suppose that we have $n \geq 1$ reports of events $\{D_y, \bar{D}_y\}$ and we wish to determine adjusted likelihood ratio Λ in either the conflicting or confirming cases. We note that when $n = 1$ we have, trivially, the confirming case.

In addition to the independence and conditional independence issues mentioned at the outset of this paper we have the following three considerations involving conditional independence.

(1) In evaluating the inferential impact of events D_y and \bar{D}_y we must ask whether or not any subset of the prior evidence in set S would condition D_y and \bar{D}_y under any hypothesis H_k . Let s be any subset of events or reports in S . Taking D_y , for example, if $P(D_y|H_k \cap s) = P(D_y|H_k)$ for every $s \in S$, then the impact of D_y on H_k is not conditioned by prior evidence. That is, D_y and S are conditionally independent given H_k . If there is some s for which $P(D_y|H_k \cap s) \neq P(D_y|H_k)$, then D_y and s are nonindependent conditional upon H_k . The problem of coping with conditional nonindependence among events has been discussed previously (Edwards, 1963; Schum, 1966; Schum, 1969; Schum, Southard, and Wombolt, 1969). Possible conditional nonindependence involving a current event and an unverified report (or reports) involves complexities that, to the authors' knowledge, have not yet been treated systematically. This problem seems worthy of consideration, however, since it is easy to imagine many instances in which one or more items of unverified prior evidence may condition a current event being evaluated.

(2) We noted that reports D^*_i and \bar{D}^*_i from source i must be conditioned by events D and \bar{D} otherwise these reports from this source can have no inferential impact. Our development of Λ in the confirming and conflicting cases assumed that $h_i = P(D^*_i|D)$ and $c_i = P(\bar{D}^*_i|\bar{D})$, for any source i , were not conditioned by any hypothesis H_k . There may be instances in which one may be advised to evaluate h_i and c_i for current reports $D^*_{y_i}$ and $\bar{D}^*_{y_i}$ in the light of some subset s of prior evidence. There appear to be two conditional independence issues.

(a) If $P(D^*_{y_i}|D_y \cap s) = P(D^*_{y_i}|D_y)$ for every $s \in S$, then h_i (hit rate on D^*_y for source i) is not conditioned by prior evidence. Similarly $P(\bar{D}^*_{y_i}|\bar{D}_y \cap s) = P(\bar{D}^*_{y_i}|\bar{D}_y)$ for every $s \in S$ means that c_i (correct rejection rate on \bar{D}^*_y for source i) is not conditioned by (i.e., is independent of) prior evidence. Coping with this conditional independence issue involves an examination of prior evidence to see whether or not there are previous inferentially relevant events or reports which would cause us to revise our estimates of the reliability parameters for any source making a current report.

(b) There may be instances in which the *joint consideration* of some hypothesis H_k and a subset s of prior evidence may cause us to revise our

estimates of the reliability parameters for some source making a current report. Formally, if $P(D^*_{y_i}|D_y \cap s \cap H_k) = P(D^*_{y_i}|D_y)$, and if $P(\bar{D}^*_{y_i}|\bar{D}_y \cap s \cap H_k) = P(\bar{D}^*_{y_i}|\bar{D}_y)$ for every $s \in S$ and for every H_k , then h_i and c_i for current reports $D^*_{y_i}$ and $\bar{D}^*_{y_i}$ are conditionally independent of the joint occurrence of any subset of prior evidence under every hypothesis. We note that conditional independence involving s or H_k by themselves does not imply conditional independence involving the joint consideration of s and H_k .

(3) In our discussions of Λ in the confirming and conflicting cases we assumed that the reports about the occurrence of D and \bar{D} from $n \geq 2$ sources were mutually independent given D and \bar{D} . That is, the sources neither decide among themselves what to report nor does any single source make his report contingent upon what other sources report. We can conceive of instances in which the *joint* reporting behavior of $n \geq 2$ sources may be conditioned by some s , some H_k , or by the joint consideration of some s and H_k .

CONCLUDING REMARKS

We have discussed multisource Λ for the case in which values of Λ are bounded by the inferential impact of events D or \bar{D} being reported by the sources. These bounds on Λ occur subject to the independence and conditional independence assumptions mentioned at the outset. The reader has surely noticed that there is a standard form of Λ which appears in every case. This standard form is:

$$\Lambda = (p_1 + v_k)/(p_2 + v_k). \quad (4)$$

In this standard form of Λ , v_k is a source reliability parameter whose value in special case k depends upon the number of sources, their individual values of A_i and B_j , and the pattern of reports made. Table 1 below summarizes values of parameter v_k and the range of values for v_k in various special cases. We emphasize that these values of v_k in Table 1 are contingent upon our independence and conditional independence assumptions. Values p_1 and p_2 in Eq. (4) are the probabilities of event D given H_1 and H_2 , respectively.

We derived statements which prescribe the inferential impact in a *joint* confirming or conflicting report from n sources, each with less than perfect reporting accuracy. Our use of the term "joint" report may not be completely desirable since it may suggest that all n reports must be available simultaneously. Using Eq. (4) and the v_k shown in Table 1 we can easily see how the inferential impact in any n reports can be assessed *sequentially* over time. We have already seen one example of this process in the discussion of "adding additional sources" in the confirming case. First suppose

TABLE 1
SUMMARY OF VALUES OF SOURCE RELIABILITY PARAMETER v_k IN
VARIOUS SPECIAL CASES

Case	Definition of v_k	Range of v
(1) Single source reports D^*	$v_1 = [A_1 - 1]^{-1}$	$v_1 \geq 0$
(2) Single source reports \bar{D}^*	$v_2 = [(1/B_1) - 1]^{-1}$	$v_2 \leq -1$
(3) All n sources report D^*	$v_3 = \left[\prod_{i=1}^n A_i - 1 \right]^{-1}$	$v_3 \geq 0$
(4) All n sources report \bar{D}^*	$v_4 = \left[\left(1 / \prod_{j=1}^n B_j \right) - 1 \right]^{-1}$	$v_4 \leq -1$
(5) Conflicting case: r sources report D^* , $n - r$ sources report \bar{D}^* .	$v_5 = \left[\left(\prod_{i \in I} A_i / \prod_{j \in J} B_j \right) - 1 \right]^{-1}$	$v_5 \geq 0$, when $\prod_{i \in I} A_i > \prod_{j \in J} B_j$. $v_5 \leq -1$, when $\prod_{i \in I} A_i < \prod_{j \in J} B_j$.

Note. We have assumed $A_i = h_i/f_i > 1$, $B_j = c_j/m_j > 1$, where: h_i = hit rate for source i , f_i = false-alarm rate for source i , c_j = correct rejection rate for source j , m_j = miss rate for source j , I = set of all r sources making reports D^* , J = set of all $n - r$ sources making reports \bar{D}^* .

that a single source, whose hit and false-alarm rates are known to us, reports D^* . We can calculate Λ using Eq. (4) with parameter v_1 . Some time later, two other sources with known hit and false-alarm rates also report D^* . We can revise Λ using Eq. (4) and v_3 with $n = 3$ since the three sources are confirming about the occurrence of event D . Now suppose that a fourth source with known correct rejection and miss rates reports \bar{D}^* . We revise Λ using Eq. (4) and v_5 with $I = 3$ and $J = 1$, since the fourth source conflicts with the other three. Adding any number of other sources merely requires a change in the value of parameter v_5 .

Our representation for Λ in the confirming and conflicting cases has greater generality than might be supposed from our developments thus far. We have only been concerned thus far with inferences involving two disjoint hypotheses H_1 and H_2 . Now suppose that our inferences concern a disjoint class $\{H_1, H_2, \dots, H_i, \dots, H_m\}$ of hypotheses which may

or may not be exhaustive. To revise our opinions about the relative likelihood of the m hypotheses on the basis of a joint conflicting or confirming reports in class $\{R_1, R_2, \dots, R_n\}$ we need only calculate the $m - 1$ independent values of Λ . This requires knowledge of p_i for every H_i and appropriate values of v_k .

REFERENCES

- EDWARDS, W. Probabilistic Information Processing in Command and Control Systems. ESD Technical Documentary Report 62-345, Electronic Systems Div., L. G. Hanscom Field, Bedford, Mass., February, 1963.
- GETTYS, C. F., & WILLKE, T. A. The application of Bayes' theorem when the true data state is unknown. *Organizational Behavior and Human Performance*, 1969, **4**, 125-141.
- PFEIFFER, P. E., & SCHUM, D. A. *Introduction to Applied Probability*, New York: Academic Press, 1973.
- SCHUM, D. A. Inferences on the basis of conditionally nonindependent data. *Journal of Experimental Psychology*, 1966, **72**, 401-409.
- SCHUM, D. A. Concerning the simulation of diagnostic systems which process complex probabilistic evidence sets. *Aerospace Medical Research Labs.* TR-69-10, April, 1969.
- SCHUM, D. A., & DU CHARME, W. M. Comments on the relationship between the impact and the reliability of evidence. *Organizational Behavior and Human Performance*, 1971, **6**, 111-131.
- SCHUM, D. A., SOUTHARD, J. F., & WOMBOLT, L. Aided Human Processing of Inconclusive Evidence in Diagnostic Systems: A summary of experimental evaluations. *Aerospace Medical Research Labs.*, TR-69-11, May 1969.