

AD-770 568

TOWARD AN ERROR THEORY FOR PIP
(PROBABILISTIC INFORMATION PROCESSING):
INFERENCE BASED ON AN ALTERNATIVE
FORMULATION OF THE DATA SPACE

Dennis G. Fryback, et al

Michigan University

Prepared for:

Office of Naval Research
Advanced Research Projects Agency

November 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

AD-770568

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 011313-9-T	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TOWARD AN ERROR THEORY FOR PIP: INFERENCE BASED ON AN ALTERNATIVE FORMULATION OF THE DATA SPACE		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER None
7. AUTHOR(s) Dennis G. Fryback and Ward Edwards		8. CONTRACT OR GRANT NUMBER(s) N00014-67-A-0181-0049
		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 197-021 ARPA Order No. 2105
9. PERFORMING ORGANIZATION NAME AND ADDRESS Engineering Psychology Laboratory University of Michigan Ann Arbor, Michigan		12. REPORT DATE 9 November 1973
		13. NUMBER OF PAGES 18 23
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
14. MONITORING AGENCY NAME AND ADDRESS (if different from Controlling Office) Engineering Psychology Programs Office of Naval Research Arlington, Virginia 22217		
16. DISTRIBUTION STATEMENT (of this Report) Approved for Public Release; Distribution Unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Probabilistic information processing Bayes's theorem Likelihood ratios Odds		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Probabilistic Information Processing (PIP) systems, as currently conceived, use experts' intuitive judgments about the diagnostic impact of individual data as inputs for mechanical aggregation by Bayes's theorem. Past research has shown that the posterior odds output by PIP are much more extreme than those arrived at via human aggregation. Because of this superior efficiency PIP-type processing of fallible data has been recommended as an important tool for decision making. The present paper questions the uncritical use in PIP of estimated likelihood ratios as if they were		

Reproduced by
**NATIONAL TECHNICAL
INFORMATION SERVICE**
U S Department of Commerce
Springfield VA 22151

TOWARD AN ERROR THEORY FOR PIP: INFERENCE BASED ON
AN ALTERNATIVE FORMULATION OF THE DATA SPACE

Technical Report

9 November 1973

Dennis G. Fryback and Ward Edwards
Engineering Psychology Laboratory
The University of Michigan
Ann Arbor, Michigan

This research was supported by the Advanced
Research Projects Agency and was monitored
by the Engineering Psychology Programs, Office
of Naval Research, under Contract Number
N00014-67-A-0181-0049, Work Unit Number
NR 197-014.

Approved for Public Release;
Distribution Unlimited



Perhaps one of the most striking findings in the area of human decision making is that of conservatism. Edwards (1968) summarizes the phenomenon:

An abundance of research has shown that human beings are conservative processors of fallible information. Such experiments compare human behavior with the outputs of Bayes's theorem, the formally optimal rule about how opinions (that is, probabilities) should be revised on the basis of new information. It turns out that opinion change is very orderly, and usually proportional to numbers calculated from Bayes's theorem—but it is insufficient in amount (pp. 17-18).

In many settings it is extremely desirable to extract from the information at hand all of the certainty that is possible, not only because it is important to make as informed a decision as possible but also because information can be very expensive. Notable areas where these considerations apply are military intelligence, medical diagnosis, business decision making, etc.

Why are men conservative information processors? The contributing factors are several (see DuCharme, 1969; and Edwards, 1968, for reviews of relevant research) but the outstanding one seems to be misaggregation of information. People seem to assess correctly the impact of a single datum, but do not combine data correctly. Edwards has proposed a probabilistic information processing system to overcome human suboptimality (Edwards, 1962; Edwards, Lindman & Phillips, 1965; Edwards, Phillips, Hays & Goodman, 1968).

In their abstract, Edwards, Phillips, Hays & Goodman (1968) summarized this system:

A Probabilistic Information Processing System (PIP) uses men and machines in a novel way to perform diagnostic information processing. Men estimate likelihood ratios for each datum and each pair of hypotheses under consideration or a sufficient subset of these pairs. A computer aggregates these estimates by means of Bayes's theorem of probability theory into a posterior distribution that reflects the impact of all available data on all hypotheses being considered. Such a system circumvents human conservatism in information processing, the inability of men to aggregate information in such a way as to modify their opinions as much as the available data justify.

Edwards, et al., go on to state their major empirical results: data which lead PIP to give 99:1 odds favoring a hypothesis resulted in 4.5:1 odds in favor of that hypothesis when the data were evaluated by a non-PIP system. They interpret this result as demonstrating that PIP used the information contained in the data much more efficiently than do competing inference systems they employed in their simulation. The implication of this superior efficiency is that in situations where many pieces of information must be processed it should be done by a PIP-like system and that system's output should be the basis of decision.



One common attribute of settings where it may be profitable to employ a PIP system for information aggregation is that it is very important to make the best decision possible. Thus the fact that a preponderance of the evidence points in one direction may not be as important as the exact odds level associated with the evidence. The difference between being 60:1 certain and being 99:1 certain may be very large in terms of the course of action that is decided upon.

In the Edwards, et al., simulation, PIP was the only processing system which reached these levels of certainty from the data. But their simulation concerned complex hypothetical scenarios of the "world of 1975" which did not allow the possibility of calculating "correct" odds. PIP was certainly non-conservative compared with the other systems, but there is no assurance that PIP may not have been too extreme, i.e., ended up more certain than the data justified. Edwards, et al., cite an experiment by Phillips (1966) which employed a PIP system in a situation where "correct" odds could be externally calculated. He found a resulting difference in processing systems comparable to the Edwards, et al., finding and furthermore reported that PIP itself appeared somewhat conservative. However, in her dissertation, Wheeler (1972) examined misaggregation as the source of conservatism. She used a data generating source for which the "correct" odds could be calculated and the results strongly point to misaggregation as the basic source of human conservatism.

She reports very conservative posterior odds associated with human-aggregated information and nearly veridical posterior odds associated with machine-aggregated information. However, in the experimental condition where the data were relatively undiagnostic the machine-aggregated posterior odds were somewhat excessive when compared with the veridical posterior odds (median machine-aggregated log odds vs. veridical log odds showed correlation of .967 and regression slope of 1.514).

But it is in exactly this sort of setting, where there is an abundance of relatively undiagnostic data, that PIP is most likely to be applied (e.g., military intelligence). It becomes evident that the development of an error theory for PIP is extremely desirable. The following discussion attempts to extend present knowledge in the direction of such a theory.

Bayes's theorem is the formally optimal tool for revising probabilistic opinion in the light of new information. Given a hypothesis and a datum bearing on that hypothesis, Bayes's theorem states that:

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)} .$$

Here, D represents the datum, H the hypothesis, $P(D|H)$ is the probability of the datum given the truth of the hypothesis, $P(H)$ the prior probability of the hypothesis, $P(D)$ the probability of observing the datum (unconditionally), and $P(H|D)$ the posterior probability of the truth of the hypothesis. For two com-

peting hypotheses, H_1 , H_2 , and a datum, D , bearing on them, Bayes's theorem may be used in odds form:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(D|H_1)P(H_1)}{P(D|H_2)P(H_2)}$$

or,

$$\Omega_2 = L \cdot \Omega_1,$$

where Ω_1 is the prior odds for H_1 against H_2 , and L is the likelihood ratio associated with D . The quantity L completely summarizes all of the diagnostic impact of D vis a vis H_1 and H_2 (see Edwards, Lindman & Savage, 1963). When a sequence of conditionally independent data are considered, Bayes's theorem may be applied iteratively, the posterior odds for the $(n - 1)$ th datum being used as the prior odds for the incorporation of the n th datum.

In a PIP system human experts estimate the relative impact of each datum, i.e., they estimate the quantity L , and a computer aggregates these according to Bayes's theorem. The computer, which is just Bayes's theorem mechanized, never observes the basic datum—it receives only an operator's estimate of the impact of the datum, the likelihood ratio. Thus to the computing system, the basic datum is the operator's estimate of the likelihood ratio.

Men seem very adept at estimating likelihood ratios (much of the literature establishing this is reviewed in Edwards, 1968). However, if error enters into the PIP output it enters at this point—where the human operators transform the information contained in the data into numbers. Even if the estimators are very good on the average in assessing the diagnostic impact of the data, their estimates are bound to be somewhat variable. That is, if on some

occasion an operator estimates the likelihood ratio for some datum to be 1.7 we cannot be sure that at some other time for that same datum and situation he might have estimated the likelihood ratio to be a little more or a little less than 1.7. Unfortunately the formal nature of probabilistic inference is such that this sort of variability doesn't just "cancel out." When additional variability is introduced into an inference system, the inevitable result is to limit the level of certainty which that system can reach. In some cases the effect of this sort of source unreliability can be quite surprising (Schum & DuCharme, 1969).

What does this mean for a PIP system which uses human estimates of likelihood ratios? When PIP receives a likelihood ratio to be entered into the computing algorithm, this ratio can be taken to mean: "a datum with this exact diagnostic value has occurred." This is the definition adopted by Edwards, et al. (1968). Given this interpretation, the report of a likelihood ratio is just as diagnostic as the observation of the datum.

A second interpretation is suggested by consideration of human variability. Here, when a likelihood ratio is given to the computer it is interpreted as "a datum has been observed and its diagnostic value is estimated to be (some number). Use this estimate to infer the true diagnosticity of the datum and use the result of that inference as input to inference algorithms for H_1 and H_2 ." Note that adopting this second interpretation introduces into the inference problem uncertainty based on the quality of the estimated likelihood ratios.

Any inference-making system which takes this datum-level uncertainty into account can never be as sure of itself (so to speak) as a system which ignores

this source of uncertainty—as does the PIP used by Edwards, et al. Maybe PIP was extracting more certainty than the data, i.e., the estimates, justify. Thus the goal of an error theory for PIP, as PIP is currently conceptualized, is to characterize the source of possible error in the system and to incorporate this characterization into the inference algorithm in a formally optimal manner. An outline for one approach to this problem comprises the rest of this paper.

Suppose that our basic problem is deciding between two competing hypotheses, H_1 and H_2 . As the basis for inference we collect data, $d_1, d_2, \dots, d_n \in D$, where D is the set of all data relevant to the two hypotheses. The diagnostic impact of a datum, d_1 , is completely described by its likelihood ratio,

$$\frac{P(d_1 | H_1)}{P(d_1 | H_2)}$$

For each d_1 there exists a $L(d_1) \in \mathcal{L}$, where

$$L(d_1) = \frac{P(d_1 | H_1)}{P(d_1 | H_2)},$$

and \mathcal{L} is the set of positive real numbers. Since the map $L: D \rightarrow \mathcal{L}$ is many:1, it makes sense to speak of a density function over the set \mathcal{L} under the function $L(D)$.

If the density function $f(d|H_1) \neq f(d|H_2)$, then it is possible to speak of two distinct density functions over \mathcal{L} , $f(L|H_1)$ and $f(L|H_2)$. Now, speaking of

the basic datum for Bayes's theorem as a likelihood ratio, it is perfectly permissible to write:

$$\frac{P(H_1|L)}{P(H_2|L)} = \frac{P(L|H_1)}{P(L|H_2)} \cdot \frac{P(H_1)}{P(H_2)}.$$

Given an L sampled from the distribution with density $f(L|H_1)$, the estimated likelihood ratio can be characterized as L' , a random variable with density $f'(L'|L)$, where the distribution of the estimated likelihood ratio is dependent only upon the value of the veridical likelihood ratio (this assumes that L' is conditionally independent of H ; i.e., $f(L'|L,H) = f(L'|L)$). Now, following the development by Gettys & Willke (1969), we may write

$$P(H_1|L') = \frac{P(L'|H_1)P(H_1)}{P(L')},$$

where

$$P(L'|H_1) = \int_{\mathcal{L}} f'(L'|L)f(L|H_1)dL.$$

(Gettys and Willke use summations; here the continuous case is considered. In either case the arguments are parallel.)

Then in odds form this becomes

$$\frac{P(H_1|L')}{P(H_2|L')} = \frac{P(H_1)}{P(H_2)} \cdot \frac{\int_{\mathcal{L}} f'(L'|L)f(L|H_1)dL}{\int_{\mathcal{L}} f'(L'|L)f(L|H_2)dL} \quad (1)$$

Taking logs, we get

$$\log \frac{P(H_1 | L')}{P(H_2 | L')} = \log \frac{P(H_1)}{P(H_2)} + \log \frac{\int_{\mathcal{L}} f'(L' | L) f(L | H_1) dL}{\int_{\mathcal{L}} f'(L' | L) f(L | H_2) dL} \quad (2)$$

The derivation of equation (1) depends on three assumptions. These may be itemized as:

- (A1) L' (the estimated likelihood ratio) is conditionally independent of H_1 and of H_2 .
- (A2) $f'(L | L_i)$ and $f'(L' | L_j)$ differ only in the value of the parameters L for all i, j .
- (A3) the value of the variable L' is known.

To actually calculate the posterior odds, given an estimated likelihood ratio, one must specify the conditional probability density functions $f(L | H_1)$, $f(L | H_2)$, and $f'(L' | L)$. The importance of the various parameters which these functions involve to the form of the ratio of integrals in equation (1) is demonstrated by the following example.

Assume that the basic datum input to the PIP system is the logarithm of a likelihood ratio estimate. Let L be the log veridical likelihood ratio, and L' be the log estimated likelihood ratio. Now, under the log transformation, \mathcal{L} is redefined as the set of all real numbers. Assume that $f(L | H_1)$ is the normal density with parameters μ and σ_h^2 . Note that by specifying $f(L | H_1)$ the density $f(L | H_2)$ is automatically specified because the relation

$$\frac{f(L | H_1)}{f(L | H_2)} = \exp(L) \quad (3)$$

must be satisfied at all points $L \in \mathcal{X}$ (here we assume that logarithms to the base e are used). Assume further that the density $f'(L'|L)$ is also the normal density function, with parameters L and σ_L^2 .

In words, it has been assumed that if H_1 were true and we collected all possible data that would distinguish between H_1 and H_2 and observed the distribution of the veridical log likelihood ratios associated with the data we would find that it was the normal probability distribution with mean μ and variance σ_h^2 . Furthermore, given a particular datum with log likelihood ratio L the estimate which would be given to the PIP system, L' , may be thought of as $L' = L + \omega$, where ω is a random error which is normally distributed with a mean of zero and variance σ_L^2 . In other words, the estimator is pretty good in that on the average when a datum of diagnosticity L occurs he will give an estimated diagnosticity $L' = L$, but may vary around the true value a little bit.

The quantity of interest is the likelihood ratio that should be used in Bayes's theorem when the estimator gives an estimate of L' . This quantity is the ratio of integrals in equation (1). Substituting into the numerator of this ratio the densities that have been assumed we get:

$$\int_{\mathcal{L}} f'(L'|L)f(L|H_1)dL = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_L^2}} \exp\left\{-\frac{1}{2}\left(\frac{L'-L}{\sigma_L}\right)^2\right\} \cdot \frac{1}{\sqrt{2\pi\sigma_h^2}} \exp\left\{-\frac{1}{2}\left(\frac{L-\mu}{\sigma_h}\right)^2\right\} dL.$$

In the denominator, using the assumptions and relationship (3), we get:

$$\int_{\mathcal{L}} f'(L'|L)f(L|H_2)dL = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \sigma_L} \exp\left\{-\frac{1}{2}\left(\frac{L' - L}{\sigma_L}\right)^2\right\} \cdot \frac{1}{\sqrt{2\pi} \sigma_h^2} \exp\left\{-\frac{1}{2}\left(\frac{L - \mu}{\sigma_h}\right)^2 - L\right\} dL .$$

Simplifying and taking logarithms we get the adjusted log likelihood ratio:

$$L'' = \log \frac{\int_{\mathcal{L}} f'(L'|L)f(L|H_1)dL}{\int_{\mathcal{L}} f'(L'|L)f(L|H_2)dL} = \frac{L'\sigma_h^2 + \mu\sigma_L^2 - \frac{\sigma_h^2\sigma_L^2}{2}}{\sigma_h^2 + \sigma_L^2} \quad (4)$$

This is the quantity that is used in (2), the log odds form of Bayes's theorem. The quantity that the Edwards, et al. (1968) study used to calculate posterior log odds would have been L' , whereas if variability in likelihood ratio estimates is taken into account the quantity that would have been used is L'' which is somewhat less than L' . Note that $L'' = L'$ if there is no possibility of error in the estimates, i.e., $\sigma_L^2 = 0$. How much less than L' is L'' depends upon the parameters σ_h , σ_L , and μ under the assumptions that have been made.

To obtain some idea of the magnitude of the adjustment from L' to L'' , consider a numerical example. In the Edwards, et al., study the posterior odds of 99:1 were achieved by the cumulative impact of 60 data. Since the particular hypothesis being considered began with odds of 1:5, the posterior odds represents a cumulative likelihood ratio of 495, or, in natural logarithms, 6.204. The average log likelihood ratio estimate was thus approximately .103. Since the model of estimator error assumes that the average estimate is the veridical value, let us assume for the sake of example that $\mu = .103$. The data reported by Edwards, et al., do not allow calculation of the variance of log likelihood ratios for the 60 data scenario leading to the posterior odds of 99:1. However,

they do give a distribution of likelihood ratios for all data used in all scenarios (some 5400 data). The overall variance may be approximated from the figures they give (page 262, Table V); using this yields the estimate $\sigma_h^2 = .0253$.

In practice, to use equation (4) one would obtain a log likelihood ratio estimate, L' , for each datum then solve for L'' , the adjusted log likelihood ratio, and use this value in equation (2) to revise the prior log odds. For the Edwards, et al., data we do not have the individual likelihood ratio estimates for the 60 data sequence. Instead, for the sake of example, we will use equation (4) to obtain L'' corresponding to an estimate of $L' = .103$, the mean for the 60 data. Multiplying the resulting L'' by 60 gives us the expected log cumulative likelihood ratio for the 60 data taken together. This value is then used to give a rough estimate of what the posterior odds might have been for Edwards, et al., data sequence if estimator variability had been taken into account. The following table displays these calculations for selected levels of σ_L^2 .

Posterior Adjusted Odds for Selected Levels of σ_L^2

σ_L^2	L''	Estimated Posterior Odds	σ_L^2	L''	Estimated Posterior Odds
.0000	.1030	99.00	.0055	.1007	84.36
.0005	.1028	92.19	.0060	.1006	83.52
.0010	.1025	93.86	.0065	.1004	82.72
.0015	.1023	92.59	.0070	.1003	81.95
.0020	.1021	91.38	.0075	.1001	81.21
.0025	.1019	90.23	.0080	.1000	80.50
.0030	.1017	89.14	.0085	.0998	79.82
.0035	.1015	88.09	.0090	.0997	79.16
.0040	.1013	87.10	.0095	.0995	78.53
.0045	.1011	86.14	.0100	.0994	77.91
.0050	.1009	85.23			

As may be seen from the calculations, the effect of estimator variability on the posterior odds may be quite slight in this situation where there is an abundance of data which is relatively slight in its diagnostic value. A reasonable question is what values of σ_L^2 one might encounter. Unfortunately the literature on likelihood ratio estimation is rarely reported in a way that allows this quantity to be estimated. Edwards, et al., do report standard deviations for repeated estimates by their subjects which may be used to estimate σ_L^2 . Their value averaged around .0007. From the table it may be seen that this amount of variability would have only minimal effect on the reported posterior odds. Wheeler (1972), whose subjects gave estimates for abstract laboratory stimuli, does not report a number comparable to σ_L^2 ; but inspection of her Figure 5, a plot of veridical log likelihood ratios against median estimated log likelihood ratios, allows us to take a rough guess that a value of .01 may not be unreasonable in the range of $\pm .4$ for the veridical log likelihood ratios. This size variability would decrease the PIP posterior odds from 99 to about 78.

It must be emphasized that the calculations that the table is based upon employ only the roughest estimates of the parameters needed for exact solutions. In addition, the model of man as an estimator that is implicit in the equations is a fairly generous one. It assumes that the only source of estimate error is random variability. This may not be the case at all—more likely, variability increases with the magnitude of the veridical log likelihood ratio (see Wheeler, 1972, Figure 5). However, as has been pointed out above, situations where PIP might be applied seem most often to be those where the data is of low diagnosticity, i.e., low average log likelihood ratio and not too much variability

around that average. The example used for the table may be considered prototypical of this type of setting.

One possible criticism of the present development concerns the distributional assumptions that have been made. The theory which culminated in equation (2) is sound regardless of the true data and estimator distributions. However, one may argue with the further development which assumes that $f(L|H)$ and $f'(L'|L)$ are Gaussian densities. We would argue that for the sake of studying analytically the effect of estimator variability on the PIP posterior odds that the forms assumed for these densities should probably serve very well as approximations in the settings being considered. Unless the data generating process is very perverse the distribution of veridical log likelihood ratios will be roughly single peaked. Once this condition is met the normal density can be considered to be a very good first approximation. If one is very concerned about the fit (or lack of it) he can always be generous in estimating the variance, since the magnitude of the adjustment in log likelihood ratio increases monotonically with variance of the normal distribution used to approximate $f(L|H)$. The result will be a conservative estimate of the adjusted posterior odds.

Assuming $f'(L'|L)$ to be the normal density may not be a bad approximation at all. This assumption is quite pervasive in the area of psychophysics. One should worry more about assuming that the variance is independent of L than assuming the f' to belong to the family of normals. Here again, though, inadequacy of fit may be compensated for by overestimating the variance which in turn overestimates the necessary adjustment to the PIP posterior odds.

The steps one would go through to apply the present error theory in an applied PIP situation follow two distinct paths. If the necessary distributions are known the theory could be applied at the level of the computing algorithm. This will almost never be the case. Assuming that the requisite distributions are unknown, one could make assumptions similar to those made here and compute, using equation (4) iteratively, datum by datum. Alternatively one could go ahead as did Edwards, et al., and use the estimates as if they were veridical; then, using assumptions we have outlined, retrospectively examine the numerical consequences of assuming various amounts of variability in the estimates as we did in the table

This latter route would be by far the easiest because it is highly likely that the exact parametric values for μ , σ_h^2 , and σ_L^2 needed for the datum by datum application can only be estimated in retrospect. Presumably the posterior odds level at which the course of action to be chosen will change is known before the data are evaluated. The retrospective application of the error theory as outlined above will show just how much error, i.e., the magnitude of σ_L^2 , can be tolerated without changing the decision. If even generous estimates of estimator variability fall within this region of tolerance, then we can feel assured that inherent variability will not be deleterious to the decision. On the other hand, if likely levels of variability will possibly change the decision, one would be advised to recompute on a datum by datum basis using the retrospective parametric estimates and use the resulting adjusted posterior odds as the basis for decision or to reserve judgment until more conclusive evidence is collected.



This last consideration, what to do if the error analysis indicates the decision could be affected, points up the major weakness of the present development. In practice there will be formidable obstacles to acquiring the necessary knowledge to apply the theory as an integral part of the aggregation algorithm. The true distributions and their necessary parameters will almost never be known making it impossible to apply equation (2) in exact form. When (2) is applied in approximate form as demonstrated here, we may be conclusive in saying when the PIP posteriors are sufficient for making a given terminal decision, but not in saying when they are not sufficient. The major advantage of the present theory is that it uses log likelihood ratio estimates as its basic input rather than individual likelihoods or probabilities, the latter task being one at which humans are far less adept (Phillips & Edwards, 1966) Other discussions of datum reliability are closely tied to this later type of judgement (see Kelly, 1972).

As a final note, it should be pointed out that the alternatives to PIP are as subject to estimator variability as is PIP. The next best system for information processing examined by the Edwards, et al., study was christened POP. In this mode, operators reestimated the actual posterior odds after seeing the data instead of estimating likelihood ratios and letting the computer aggregate for them. When PIP was reporting odds of 99:1, POP was giving 4.5:1. But again, if we were to take into account inherent variability in the POP posterior odds estimates, the 4.5:1 odds would be reduced as are the PIP odds. The same basic approach used here would lead to a similar error theory for POP, however some

of the algebra would be different. Such a development is beyond the scope of the present paper.

Where to go from here? The theory that has been developed is sufficiently promising to warrant more intensive work. The goals of further research should include both extension and refinement of the theory and broadening the empirical basis for application. The first obvious extension of the theory is obtaining concise formulas for adjusted likelihood ratios such as equation (4), but basing the derivation on distributional assumptions other than the ones used here. Once concise forms are obtained, there should be numerical work done to examine the magnitude of error possibly introduced by using the wrong assumptions as approximations. We expect that most reasonable distributions will not significantly change the conclusions we have reached using the normal approximations; but it would be nice to have the numerical work to support this statement.

For empirical work, the first step would be to characterize the distributions of data likelihoods that may be encountered in application of PIP in a variety of information processing settings. A concurrent line of investigation would comprise building an empirical data base from which the $f'(L'|L)$ distributions could be obtained. One byproduct of this would be a characterization of those inference problems in which estimator variability is sufficient to warrant applying the theory on a datum by datum basis rather than the retrospective fashion we have described.

References

- DuCharme, W. M. A review and analysis of the phenomenon of conservatism in human inference. Rice University, Interdisciplinary Program in Applied Mathematics and Systems Theory; Systems #46-5. Dec., 1969.
- Edwards, W. Dynamic decision theory and probabilistic information processing. Human Factors, 1962, 4, 59-73.
- Edwards, W. Conservatism in human information processing. In B. Kleinmuntz (Ed.), Formal Representation of Human Judgment, New York: Wiley, 1968, 17-52.
- Edwards, W., Lindman, H. & Phillips, L. D. Emerging technologies for making decisions. In New Directions in Psychology II. New York: Holt, Rinehart, and Winston, 1965, 261-325.
- Edwards, W., Lindman, H. & Savage, L. J. Bayesian statistical inference for psychological research. Psychological Review, 1963, 70, 193-242.
- Edwards, W., Phillips, L. D., Hays, W. L. & Goodman, B. C. Probabilistic information processing systems: Design and evaluation. IEEE Transactions on Systems Science and Cybernetics, 1968, SSC-4, 248-265.
- Gettys, C. F. & Willke, T. A. The application of Bayes's theorem when the true data state is uncertain. Organizational Behavior and Human Performance, 1969, 4, 125-141.
- Kelly, C. W. Application of Bayesian procedures to hierarchical inferences. Doctoral Dissertation, The University of Michigan, 1972.
- Phillips, L. D. Some components of probabilistic inference. Doctoral Dissertation, The University of Michigan, 1966.
- Phillips, L. D. & Edwards, W. Conservatism in a simple probability inference task. Journal of Experimental Psychology, 1966, 346-354.
- Schum, D. A. & DuCharme, W. M. Comments of the relationship between the impact and reliability of evidence. Rice University, Interdisciplinary Program in Applied Mathematics and Systems Theory, Systems 46-8, August, 1969.
- Wheeler, G. F. Misaggregation versus response bias as explanations for conservative inference. Doctoral Dissertation, The University of Michigan, 1972.

Distribution List

Director,
Engineering Psychology Programs
Code 455
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

(5 cys)

Office of the Chief of Naval
Operations, Op-095
Department of the Navy
Washington, D. C. 20350

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314

(12 cys)

Dr. John J. Collins
Office of the Chief of Naval
Operations, Op-987F
Department of the Navy
Washington, D. C. 20350

Director, ONR Branch Office
Attn: Dr. C. Harsh
495 Summer Street
Boston, Massachusetts 02210

CDR H. J. Connery
Office of the Chief of Naval
Operations, Op-987M4
Department of the Navy
Washington, D. C. 20350

Director, ONR Branch Office
Attn: Dr. M. Bertin
536 S. Clark Street
Chicago, Illinois 60605

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code AX
Washington, D. C. 20380

Director, ONR Branch Office
Attn: Dr. E. Gloye
1030 East Green Street
Pasadena, California 91106

Mr. John Hill
Naval Research Laboratory
Code 5634
Washington, D. C. 20375

Director, ONR Branch Office
Attn: Mr. R. Lawson
1030 East Green Street
Pasadena, California 91106

Director, Naval Research Laboratory (6 cys)
Technical Information Division
Code 2027
Washington, D. C. 20375

Office of Naval Research
Mathematical Sciences Division
Code 434
Department of the Navy
Arlington, Virginia 22217

Director, Naval Research Laboratory (6 cys)
Attn: Library, Code 2029 (ONRL)
Washington, D. C. 20375

Office of Naval Research
Code 437
800 North Quincy Street
Arlington, Virginia 22217

Office of Naval Research
Code 463
800 North Quincy Street
Arlington, Virginia 22217