ARI Research Note 95-17

People Misinterpret Conditional Probabilities

Robert M. Hamm
University of Colorado

Research and Advanced Concepts Office
Michael Drillings, Acting Director

January 1995

ARMY RESEARCH INSTITUTE

United States Army
Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract for the Department of the Army

University of Colorado

Technical review by

Michael Drillings

Accesion For		
NTIS CRA&I		X
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1995, January	Final Sep 86 - Feb 89

4. TITLE AND SUBTITLE
People Misinterpret Conditional Probabilities

5. FUNDING NUMBERS
MDA903-86-K-0265
01102A
20161102B74F
TA BR45
WU C70

6. AUTHOR(S)

Hamm, Robert M. (University of Colorado)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Colorado at Boulder
Institute of Cognitive Science
Box 345
Boulder, CO 80309-0345

8. PERFORMING ORGANIZATION REPORT NUMBER
PN 89-4

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
ATTN: PERI-BR
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARI Research Note 95-17

11. SUPPLEMENTARY NOTES
Contracting Officer's Representative, Michael Drillings.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT *(Maximum 200 words)*

This final report summarizes the results of the project "The Use of Protocol Analysis and Process Tracing Techniques to Investigate Probabilistic Inference." In probabilistic inference, people use uncertain information to change uncertain beliefs. That is, they must integrate base rate information (about what usually happens) with uncertain information about what is happening in the present. The research shows that the most recently presented information is given undue attention. Further, although subjects recognize that the base rate information in probabilistic inference word problems is relevant, they do not give it enough impact in their considerations. This is not because of their tendency to use available numerical expressions of probability as their response, but because of their inability to interpret conditional probabilities appropriately. Specifically, the subjects think that the conditional probability p (evidence/hypothesis), which is given in the word problems and which should be taken as an input to Bayes' Theorem, is p (hypothesis/evidence), which is the output of Bayes' Theorem and which is the answer that they are asked to produce. This mistake causes subjects to produce answers that are independent of base rate information.

14. SUBJECT TERMS
Probabilistic information
Bayes' Theorem
Base rate information

15. NUMBER OF PAGES
23

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Robert M. Hamm. University of Colorado
ARI Contract MDS-903-86-K-0265

Abstract.

The results of the project "The Use of Protocol Analysis and Process Tracing Techniques to Investigate Probabilistic Inference" are summarized in this final report. In probabilistic inference, people use uncertain information to change uncertain beliefs. That is, they must integrate base rate information (about what usually happens) with uncertain information about what is happening in the present case. It was shown that the most recently presented information is given undue attention. Further. although subjects recognize that the base rate information in probabilistic inference word problems is relevant. they do not give it enough of an impact in their considerations. This is not due solely to their tendency to use available numerical expressions of probability as their response. Rather, it is due to their inability to interpret conditional probabilities appropriately. Specifically, the subjects think that the conditional probability p(evidence/hypothesis), which is given in the word problems and which should be taken as an input to Bayes' Theorem, is p(hypothesis/evidence), which is the output of Bayes' Theorem and which is the answer that they are asked to produce. This mistake causes subjects to produce answers that are independent of the base rate information, although they believe that the base rate information should have an impact on their answer.

People Misinterpret Conditional Probabilities: Final Report of Project Using Protocol Analysis and Process Tracing Techniques to Investigate Probabilistic Inference.

Executive Summary.

SCIENTIFIC OBJECTIVES: Taking correct action in uncertain conditions requires the ability to use two types of information: statistical base rate information about what is **likely to** happen, based on the history of what usually happens, and imperfectly reliable information about what is **now** happening. Past work on probabilistic inference has demonstrated that when these two types of information conflict, people tend to neglect the base rate information and to put unwarranted confidence in the information about the present situation, even though this information is unreliable. The goal of this research is to describe the process by which novices, as well as experts in probabilistic inference and experts in the substance of the problem, combine the two types of information in making probabilistic inferences. Understanding of both novice and successful inference processes will help us to teach strategies for correct reasoning, and design information environments that support accurate inference.

APPROACH. It is assumed that in making probabilistic inferences people use strategies that involve (a) understanding the information given in the problem, and (b) working with it to estimate the probability. To investigate these strategies, subjects are asked to solve probabilistic inference word problems. Their answers, their choices of information, and their concurrent verbalizations are analyzed for evidence concerning how they interpret the given information and what they do with it.

FINDINGS. In Study I, novices were asked to estimate the probability that a hypothesis was true in three probabilistic inference word problems. In each problem, they answered before and after the presentation of each of three types of information -- base rate, evidence, and reliability of evidence. Many subjects responded with numbers that were available in the problem presentation. The more recent information had a greater impact. Comparison of production system simulations of the typical responses and the normative responses suggested that the neglect of the base rate information was due in part to a misunderstanding of the reliability information. Specifically, subjects did not distinguish between two conditional probabilities, the probability that particular evidence would be seen if a hypothesis were true p(e/h), and the probability that a particular hypothesis would be true if evidence were seen p(h/e).

Study II investigated whether the neglect of base rate could be due to an artifact of the experimental method: the fact that probilities are presented as numbers rather than verbal expressions, which may induce subjects to respond using available numbers. Although verbal presentation of probabilities and verbal responses reduced the level of use of available probabilities as the response, subjects still neglected base rate on the verbal probability problems, and their responses were no more accurate with the verbal probabilities than with numerical probabilities.

Study III focussed on the hypothesis that subjects neglect base rate because they confuse the conditional probabilities p(e/h) and p(h/e). One part of this study tested whether subjects respond any differently when presented with p(h/e) information instead of the p(e/h) usually used in these studies. There was little difference. In addition, analysis of subjects' verbalizations while considering the conditional probability information revealed that the conditional probability presented had little influence on the conditional probability concept the subject used. Finally, analysis of subjects' preferences for order in which to receive information on these problems indicates that they value base rate

information on a par with evidence information, and that they do not make a sharp distinction between the conditional probabilities. Generally, the hypothesis that subjects do not correctly distinguish p(e/h) and p(h/e) in probabilistic inference was supported, while there was little support for the hypothesis that they think that base rate information is irrelevant.

CONTRIBUTIONS TO BASIC SCIENCE. The technique of requiring answers after each possible subset of the probabilistic inference word problem information has allowed unequivocal elimination of the hypothesis that subjects completely ignore base rate. It also provides data for testing production system models of inference strategies. The discovery that people respond using available numerical probabilities, but not verbal probabilities, introduces a new dimension into discussions of the generality of flaws in people's statistical reasoning. The application of verbal protocol analysis and process tracing to the problem of probabilistic inference has strongly supported the theory that people can not appropriately distinguish the conditional probabilities p(e/h) and p(h/e), which in turn accounts for the apparent neglect of base rate that has previously been observed with probabilistic inference word problems.

POTENTIAL APPLICATIONS. Many military operational contexts require the integration of information about expectancies (prior probabilities that a hypothesis will be true) with uncertain information about what is happening at present. If the statistical information is neglected, it could lead to an excessive amount of "false alarms". If, as demonstrated here, the most recent information is given more attention, then the flow of information in operational situations should be designed so that base rate information is presented concurrently with or after the current information, so that it is not neglected. If novices or experts have difficulty distinguishing the two types of conditional probability information, then training should be designed to overcome this difficulty. Further, requests for probabilistic inferences in operational contexts should be couched in terms that do not present the opportunity for a misinterpretation of conditional probabilities to lead someone to neglect base rate information. The demonstration that verbal probabilities do not produce generally worse performance than numerical probabilities suggests there may be a legitimate role for verbal expressions of uncertainty in operational contexts.

People Misinterpret Conditional Probabilities: Final Report of Project Using Protocol Analysis and Process Tracing Techniques to Investigate Probabilistic Inference.

1. Introduction.

Probabilistic reasoning is a basis for action in a wide variety of vital contexts. A decision maker in a combat situation must interpret potentially unreliable intelligence information concerning enemy troup movements. An officer must draw conclusions concerning a new subordinate given stereotypical expectations based on the subordinate's appearance, sex, ethnicity, or way of speaking, and on impressions derived from brief interactions with the subordinate. It is important to understand how people make probabilistic inferences, what determines their accuracy or inaccuracy, and how their accuracy can be improved. Erroneous methods of interpretation of battlefield intelligence could unnecessarily decrease the probability of victory. Methods of evaluating subordinates that do not take account of both reasonable expectations based on the subordinate's group, and evidence about the individual, could lead to inefficient allocation of manpower resources, as well as resentment and low morale.

A story related by Venn (1888, cited by Niiniluoto, 1981) illustrates the difficulty that probabilistic inference poses for people. Your friend tells you that a particular number has won a lottery of 10,000 tickets. You check your pocket and discover that you hold the ticket with that number. This friend has been proven in everyday situations to recall numbers correctly 99% of the time. What is the chance that your ticket is the winner? According to Venn, the common person does not know whether the correct answer is 1 in 10,000 (the probability of the lottery, i.e., the "base rate") or 99 in 100 (the witness's reliability, p(evidence/hypothesis)). The person educated in formal probabilistic inference knows that both the base rate and the imperfectly reliable evidence are pertinent, and the answer is approximately 1 in 102 (by Bayes' Theorem).

$$p(H/E) = \frac{p(E/H)*p(H)}{p(E/H)*p(H) + p(E/\neg H)*p(\neg H)}$$

The problem of probabilistic inference, that is, the integration of prior expectations and unreliable evidence, is unavoidable whenever the status of a situation is not known for certain and information about it might be unreliable. The availability of decision support systems designed to help with probabilistic inference does not mean that people no longer need to understand the role of both types of information -- prior expectations and imperfectly reliable evidence -- in inference. Rather, decision support systems add a new level of complexity to the problem of probabilistic inference. Two incidents involving the U.S. Navy in the Persion Gulf illustrate that understanding of probabilistic inference is needed for the successful use of decision support systems.

In the case of the Stark, there was a defensive computer system that evaluated evidence that might indicate a threat from the air. This system had been producing so many false alarms, i.e., inferences of an attack when there really was none, that it had been shut off. Therefore when an Iraqi missile was fired at the Stark, it was not detected in time. A more sophisticated decision support system, operated by people with a full understanding of probabilistic inference, might have provided the commander with the capability of adjusting the warning threshold in response to varying probabilities of attack, so that it would not have been necessary to shut down the warning system completely.

The case of the Vincennes represents the opposite kind of mistake -- hitting a non-attacker instead of missing an attacker. The captain believed the word of a junior officer responsible for reading a radar, who mistakenly reported that the plane in question was descending (consistent with an attack) rather than ascending (consistent with the usual flight pattern of an Iranian civilian airliner). If there was a failure of decision making here, it was that in the 4 1/2 minute period during which the decision was made, no one thought to check the junior officer's reading of the radar, considering the possibility that it might have been less reliable than usual under the stressful conditions on the bridge (Fogarty, 1988).

In both unaided and aided decision making, people are required to make correct inferences under uncertainty. Past psychological research has shown that people's use of probabilistic information in reasoning deviates from the uses that are prescribed by the normative methods of probabilistic inference. For example, neglect of base rate information (Bar-Hillel, 1980; Tversky and Kahneman, 1982) and of the possibility of a false alarm (Doherty, Mynatt, Tweney, and Schiavo, 1979) have been shown in a number of probabilistic inference word problem studies.

The primary goal of this project has been to develop an understanding of the variety of strategies that people can use to make probabilistic inferences, so that we can know how people can do this reasoning most accurately. Discovery of accurate heuristic strategies that leaders and decision makers could be taught to use as a mental habit, as part of their automatic interpretation of the world, could lead to accurate performance of probabilistic inference in uncertain situations without reliance on external computer aids, such as those which perform Bayes' Theorem calculations. These aids have had low acceptance in decision making contexts (cf Shortliffe, 1984), partly because of fear-based psychological barriers in potential users, partly because of the practical inconvenience of the requirement of accurately entering the full set of pertinent data in the system, and partly because of the potential for catastrophic results due to minor "clerical" errors (Hammond, 1981; Hamm, 1988b; but see MacGregor, Lichtenstein, and Slovic, 1988). Methods of probabilistic inference that are well-founded, even if not perfectly accurate, and that can be integrated into decision making practice, may potentially be of great value.

The project of discovering such aids must be based on a realistic understanding of people's current capabilities of probabilistic inference, and of the processes they use in making such inferences.

2. Previous research on probabilistic inference.

Two bodies of research have studied people's unaided probabilistic inference. The first is the book-bag and poker chip paradigm (reviewed by Edwards, 1968, and Slovic and Lichtenstein, 1971), which investigated peoples' combination of information about a prior probability (summary of beliefs before new observations) plus multiple competing unreliable observations. This approach compared people's performance to the odds-likelihood form of Bayes' Theorem. The second approach is the diagnostic word problem paradigm (reviewed by Tversky and Kahneman, 1982), which investigated people's combination of a prior probability with a single unreliable observation that is inconsistent with the prior expectation. This approach compared people's performance to the simple form of

Bayes' Theorem. While some have concluded that the book-bag and poker-chip approach compares people with too difficult a standard (von Winterfeldt and Edwards, 1986), the diagnostic word problems of the second approach are similar to situations in which most people occasionally find themselves, and with which some experts must deal on a regular basis, with important consequences (e.g., medical doctors; see Eddy, 1982; and intelligence analysts; see Cohen et al, 1985, and Schum, 1987). Therefore, the present work adopts the latter approach.

A typical probabilistic inference word problem is the Cab problem, used by Tversky and Kahneman (1982) and others. The Cab problem tells subjects that

"A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85% of the cabs in the city are Green and 15% are Blue.
(b) A witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.

What is the probability that the cab involved in the accident was Blue rather than Green?" (Tversky and Kahneman, 1982, pp 156-157).

Research on probabilistic inference word problems has almost universally found that people's numerical answers differ from those produced by applying Bayes' Theorem. Even when subjects' median answer is very accurate, researchers have concluded that people are not actually calculating Bayes' Theorem, and are producing "fairly close to optimal answers for the 'wrong' reasons" (Ofir, 1988, p 361).

Researchers have emphasized two features of people's probabilistic inferences. Bar-Hillel (1980), Kahneman and Tversky (1972; Tversky and Kahneman, 1982), and others have argued that people neglect the base rate information. Doherty, Mynatt, Tweney, and Schiavo (1979) and Beyth-Marom and Fischhoff (1983) have said that people underutilize or ignore the false alarm information, $p(e/\neg h)$, the probability that the evidence that favors hypothesis h could have been seen if h were not true. Let us consider these findings in detail.

2.0.1. The neglect of base rate.

Kahneman and Tversky (1972), Lyon and Slovic (1976), Bar-Hillel (1980), and many others have shown that on word problems with a low base rate [for example, in the Cab problem the base rate $p(h)$ is that only 15% of the cabs in the city are blue], a moderate reliability [e.g., the probability of calling a blue cab "blue," $p(e/h)$, is .80], and the complementary false alarm rate [the probability of calling a green cab "blue," $p(e/\neg h)$, is .20], subjects' answers are too high. Where the answer calculated for these particular values using Bayes' Theorem is .41, subjects' median and modal answer is .80 (Kahneman and Tversky, 1972; Bar-Hillel, 1980). This does not mean, however, that people do not believe that base rate information is pertinent.

Several studies have shown that people's answers on these problems vary in response to base rate. When only the base rate was presented, without evidence, people used it as their answer (Lyon and Slovic, 1976; Hamm, 1987a; Ofir, 1988). When the base rate was made to seem causally connected to the present case, people paid more attention to it (Bar-Hillel, 1980). Studies which varied the level of base rate in repeated presentations of the same problem showed that people respond to it (Fischhoff, Slovic and Lichtenstein, 1979; Birnbaum and Mellers, 1983; Ofir and Lynch, 1984). However, Fischhoff and Bar-Hillel (1984) cautioned that in the repeated presentation studies the experimental situation may have demanded just such a response: what else is there for the subject to respond to, in repeated versions of the same problem, but the numbers that are varying in the problems? In reply, Ofir and Lynch (1984) and Ofir (1988) varied the base rate in between-subjects designs and found that the mean and median answers were usually responsive to base rate differences. Ofir (1988) found that base rate was attended over the range of possible hit rates p(e/h) and false alarm rates p(e/¬h), except when hit rate was high and false alarm rate low.

Thus only when the hit rate and false alarm rate both support the evidence, by being high and low, respectively, is competing base rate information neglected. The problems studied by Kahneman and Tversky (1972), Bar-Hillel (1980), and others have just these characteristics. Although these problems represent only a subset of the possible probabilistic inference word problem variants, it is an important subset. These are situations where the evidence is inconsistent with prior expectations. It remains to be explained why people attend the evidence in probabilistic inference word problems, while they may be governed by their expectations or prejudices in other situations. (Attributing it to the salience or perceived relevance of the base rate is not satisfactory because it is nearly tautological.)

2.0.2. The neglect of false alarm information.

To determine whether people neglect information concerning the possibility that the alternative hypothesis is true and has been mis-identified, Ofir (1988) presented problems with a range of p(e/¬h) values and found that in most cases subjects' answers vary in the appropriate direction in response to these variations. The exceptions are that false alarm information seems to be ignored either when the false alarm rate is very low, or when both the base rate and the hit rate are very high. These qualitative inconsistencies with Bayes' Theorem (Principle 4) seem to be due to simplifying strategies -- completely ignoring variables that have a small but important impact. As such, these deviations do not require special explanation. Ofir (1988) additionally noted, following Beyth-Marom and Fischhoff (1983), that while people know how to use the false alarm information if they have it, they do not think to look for it if it has not been given to them.

Our review of the qualitative features of people's performance on probabilistic inference word problems has shown that (a) although in most situations their responses covary appropriately with changes in the key base rate and reliability information, people do not seem to be integrating the cues, particularly when the cues have competing implications; (b) people's numerical responses are sometimes quite far from the Bayes' Theorem answer even when they covary appropriately with all inputs (see figures in Ofir, 1988). This means that those who must submit their health or security to decision systems that rely on people's intuitive probabilistic inference have something to worry about. And those of us who venture to offer advice on these decision systems need to understand what people are doing on these tasks.

3. Research studies in the current project.

Three studies were done. The first required subjects to estimate the probability that a hypothesis is true, after every possible combination of the three pieces of information (base rate, evidence, and reliability of evidence) usually given in probabilistic inference word problems. The second study substituted verbal expressions of probability for the usual numerical expressions to see whether the most common strategies are also used with verbal probabilities. The third study used process tracing, protocol analysis, and memory recall techniques to investigate the theory that people can not distinguish between the conditional probabilities p(e/h) [the probability that particular evidence will be seen if a hypothesis is true] and p(h/e) [the probability that a hypothesis is true if particular evidence is seen].

3.1. Study I. Subjects' estimates of p(h) following every possible subset of the key information.

A questionnaire study in which 265 undergraduate students answered three probabilistic inference problems has been completed. Four papers are based on the data from this study.

3.1.1. Basic results.

Extensive analyses of the results of the Questionnaire Study are reported in the paper "Diagnostic Inference: People's Use of Information in Incomplete Bayesian Word Problems" (Hamm, 1987). The word problems were the Cab problem (see above) and two variants, the Doctor problem and the Twins problem. The procedure in Study I differed from the procedure in the Cab problem example, however, in that subjects were required to respond with their probability that the named hypothesis is true four times during the problem, rather than just one: after the basic situation is described and again after each piece of key information is presented. The three pieces of key information are the evidence *e* (e.g., in the Cab problem, that the witness reported a Blue cab), the reliability of the evidence *p(e/h)* (that the witness correctly identifies a given cab color 80% of the time) and the base rate *p(h)* (that 15% of the cabs in the city are Blue). The three pieces of information were presented in each possible order, to different subjects. This allows us to study how subjects make probabilistic inferences when given every possible subset of the three pieces of information, e.g., when presented with only the evidence and the base rate.

Three classes of hypothesis were proposed to explain how people answer these word problems and why the answers often seem to neglect the base rate information -- 2 variants of normative probabilistic reasoning, 5 types of heuristic strategy, and 13 variants of non-normative information integration. Findings included:

1. Many subjects responded with numbers that are available in the problem presentation. Often the use of an available number is normatively correct, which implies that the novices have some understanding about appropriate reasoning in these situations. However, many of the subjects' wrong answers also used numbers available in the word problem. This implies that they may be adopting the simple strategy of answering with whatever numbers are available. It is therefore possible that the subjects who answered correctly may have done so just by luck.

2. The more recent information has a greater impact. For example, when the subjects had all three pieces of information, if base rate information was presented most recently, more subjects took it into account in producing their answers than if it had been presented first and the evidence and the reliability information had followed it. This identifies another condition that influences the subjects' likelihood of using or neglecting the base rate information (see Bar-Hillel, 1980).

3. There is no universally applied weighted averaging scheme that accounts for the average response in all conditions. Rather, some form of "contingent strategies" theory is needed to account for the data. "Contingent strategies" means, broadly, that people will adopt different information processing strategies in different conditions (when given different combinations of information), rather than applying one strategy (weighted averages) in all conditions.

4. The neglect of the base rate information is due in part to a misunderstanding of the reliability information, specifically, a confusion between p(h/e) and p(e/h) -- e.g., "the probability that it really was a blue cab if the witness called it 'blue'" and "the probability that the witness would call it 'blue' if it really were blue."

3.1.2. Rule-based models of possible response strategies.

In the paper "Explanations of the use of reliability information as the response in probabilistic inference word problems" (Hamm, 1987b), three competing hypotheses about how subjects respond to probabilistic inference word problems were described. These are (a) that subjects consider the base rate[1] to be irrelevant in principle, (b) that subjects interpolate between the base rate probability and 1.0, and then select their response from among nearby numbers that are available in the word problem, and (c) that subjects confuse the conditional probability p(e/h),[2] which is given in the problem and is an appropriate input into a Bayes' Theorem integration, with the conditional probability p(h/e)[3], which is the output of Bayes' Theorem and is an appropriate answer to the problem. These theories were expressed in production system models that represent each theory as a set of contingent strategies. It was found that subjects' answers on the subsets of the possible information were not able to narrow the hypotheses. That is, it was possible to make models, consistent with each of the three theories, that exactly predicted the most common response made by subjects in each of the possible situations (situations are defined by combinations of available information).

Despite the fact that subjects' answers were consistent with each of the three theories, the exercise of specifying them was judged to be useful. In particular, the theory that people confuse the conditional probabilities p(e/h) and p(h/e) has the potential of casting a new perspective on the "neglect of base rate". No rule in the production system model expressed a process that would be characterized as underweighting base rate. Rather, when reliability, evidence, and base rate information were all present, the rules took the reliability information p(e/h) to be p(h/e), which is the answer the problem asks for, and hence used it as the answer. The subject's answer, which seems independent of base rate, may be a reasonable response given the interpretation of p(e/h) as p(h/e), rather than being the result of a mistaken judgment of the relative relevance of statistical (base rate) and case (evidence) information, as in the view of Bar-Hillel (1980).

As just described, this study has called attention to a potential barrier to accurate probabilistic inference, which is that people may not know how to interpret the conditional probabilities in which the reliability information is often couched in probabilistic inference word problems. Training would presumably correct this problem. Even if this barrier were to be surmounted, we still lack knowledge of how to train people to best integrate the statistical and case information (but see Lichtenstein and McGregor, 1984). Yet no progress at all can be possible when subjects confuse p(h/e) and p(e/h).

3.1.3. An alternative to contingent strategies models of probabilistic inference.

The production systems described above and in Hamm (1987b) represent the information processing or artificial intelligence school of modeling cognition. A distinct approach is to model subjects' behavior as involving intuitive judgment and choice processes. For example, subjects' responses could be produced by a two stage process,

1. an intuitive judgment of the probability of the hypothesis,

---

[1] E.g., the proportion of blue cabs in the city.

[2] E.g., the probability the witness would call a blue cab "blue".

[3] E.g., the probability that the cab involved in the accident was truly blue if the witness identified it as "blue".

2. a probabilistic choice process which selects one of the available numbers as the answer, as a function of how near it is to the intuitive judgment.

The paper "A model of answer choice on probabilistic inference word problems" (Hamm, 1987c) compares these two models in terms of their assumptions and the ease with which they account for 5 aspects of the data from the Questionnaire Study. It suggests methods for combining the advantages of the two approaches.

3.1.4. Complementarity and Resuscitation.

Two additional questions can be addressed using the data from Study I. These are the subjects' understanding of the complement of a probability, and the occurrence of "resuscitations", i.e., judging the probability of a hypothesis to be 0 at one stage, and then to be a non-zero number at a later stage.

Complementarity. The questionnaire asked subjects not only for the probability that a particular hypothesis was true (e.g., that the cab involved in the accident was Blue) but also for the probability of the complementary hypothesis (that the cab was Green). Given the word problem's definition of these events as mutually exclusive and exhaustive, the correct answer to the second question is the probabilistic complement of the first, i.e., p(Green) = 1 - p(Blue). It was found that a high proportion of subjects gave complementary answers. Evidence was sought for subjects' use of variant conceptions of subjective probability, such as one proposed by Schafer (1976) in which someone with little evidence might have very low subjective probability for both a hypothesis and its complement (see Kahneman and Tversky, 1982), so that the probabilities would add up to much less than one. If this theory is correct, then as the subjects get more information, the sum of their probabilities for the mutually exclusive and exhaustive events should approach 1.0. This pattern occurred very rarely among the subjects whose answers were noncomplementary.

Resuscitations. If a Bayesian probability estimator is receiving a stream of information pertinent to the estimate of the probability of a hypothesis, and if the probability ever hits 0 or 1.0, there is no way that it can return to an intermediate value. Subjects' probabilities for a hypothesis have been observed to be "resuscitated" after hitting 0, and also to return from 1.0 (Schum and Martin, 1980; Robinson and Hastie, 1985). Such behavior was observed in this study, as well, though it was infrequent.

The import of our analysis of complementarity and resuscitations is that most naive subjects follow these rules of probability. This finding contradicts some pessimistic conclusions reached on the basis of previous research, concerning people's general inability to do any type of probabilistic reasoning. However, the occasional occurrence of noncomplementary estimates and of resuscitations should alert us to the possibility that people use numerical probabilities to mean something other than what a strict interpretation of the numbers would imply (Kahneman and Tversky, 1982).

3.2. Study II. Verbal versus numerical expressions of the probabilities in the word problems.

The second study involved using verbal rather than numerical expressions of probability in the probabilistic inference word problems, either for presentation of the stimulus probabilities or for the subjects' response probabilities. Two papers have resulted from this study, one covering the subjects' probabilistic inference using verbal probabilities, and the other exploring a methodological issue encountered when using verbal probabilities.

3.2.1. The neglect of base rate in probabilistic inference word problems occurs with verbal expressions of probability.

In Study I it was demonstrated that subjects use available numbers as their responses in probabilistic inference word problems. It is possible that they are unfamiliar with inference using numerical probabilities, and think that they are supposed to make use of the available numbers. This may account for the relative neglect of base rate information. In contrast, people usually encounter probabilistic inference in situations where the probabilities are not formally measured. In such situations, their responses may have a different character.

To test this possibility, variants of standard probabilistic inference word problems were produced that had verbal expressions rather than numerical expressions of probability. The results were reported in "Accuracy of probabilistic inference using verbal *versus* numerical probabilities" (Hamm, 1988a). Subjects received either the verbal or the numerical versions of the word problems. Thus, the baserate in the Cab problem would be presented in the sentence "When one sees a cab in the streets of the city, it a Blue cab only 15% of the time" in the numerical presentation version, and in the sentence "When one sees a cab in the streets of the city, it is seldom a Blue cab" in the verbal presentation version.

In addition, a technique for eliciting verbal responses from a given set of expressions was developed. Subjects were required to select one of 19 phrases from a list of phrases that expressed a range of probabilities from "absolutely certain" to "absolutely impossible" in steps of approximately .05. Subjects were randomly assigned to 4 conditions, verbal or numerical probabilities in the word problems crossed with verbal or numerical responses: N-N, N-V, V-N, and V-V.

Did subjects used available expressions of probability as often when the stimulus probabilities and responses were verbal, as when they were both numerical? Comparing the N-N and V-V conditions showed that subjects used available probabilities more often when they were numerical than when they were verbal.

Was this tendency to use available probabilities, in essence an artifact of the presentation of word problems which require people to deal with numerical probabilities, responsible for the neglect of base rate that had been previously observed? There was still a substantial neglect of base rate with verbal probabilities, and people's performance was not significantly worse when they were given information, and responded, using numerical probabilities.

3.2.2. The method of selecting a verbal expression of probability from a list.

A new method for eliciting judgments of probability using verbal expressions of probability was used in this study -- presenting a range of verbal expressions of probability and requiring the subject to select the most appropriate expression. An auxiliary technique is to have the subject subsequently assign numerical probabilities to each verbal expression, to facilitate interpretation. To allow us to determine whether the order in which the verbal expressions were presented affects their use, we presented the verbal expressions in four lists -- two ordered (ascending or descending) and two random. The results are discussed in "Evaluation of a method of verbally expressing degree of belief by selecting phrases from a list" (Hamm, 1988c). It was found that when the verbal expressions were arranged in a *random* order, the ordinal position of an expression in the list had a very minor, but statistically significant, effect on the selection of expressions -- people tended to select expressions that appeared in the second half of the list. This position effect was not significant with *ordered* lists, so ordered lists are recommended. Considerations of accuracy and interpersonal agreement also support the use of ordered lists of verbal expressions of probability.

Presenting a list of verbal expressions of probability is a method that can be used with respondents who are not comfortable with or capable of using numerical expressions of probability, as long as the list of expressions covers the required range with sufficient density, and as long as high precision is not required. It might also be helpful to present the verbal expressions along with commonly accepted numerical interpretations, to allow the advantage of verbal expression (subject familiarity) while countering the disadvantage (unspecified meaning).

3.3. Study III. Tests of the hypothesis that people confuse conditional probabilities.

The third study focussed on the comparison of the three hypotheses that were studied in the contingent strategies paper (Hamm, 1987b). The main new tool used here was verbal protocol analysis. In addition, subjects were given the opportunity to select the order in which to receive the base rate, evidence, and reliability information, in order to trace their information seeking, and they were required to recall a problem later, to show what concepts they had retained.

3.3.1. Pilot study of thinking aloud.

A pilot study was done (with Edson Sellers) to determine the feasibility of coding the transcripts of subjects' verbalizations while solving probabilistic inference word problems. Since this will not be written up elsewhere it is described here. Ten student subjects thought aloud while solving two or three word problems: variants of the Cab problem, the Twins problem, and/or the Doctor problem (see Hamm, 1987a). Their answers at each juncture in the problem were transcribed, unitized into sentences, and coded with respect to whether they mentioned the base rate $p(h)$, the probability of the complementary hypothesis $p(\bar{h})$, the reliability $p(e/h)$, the likelihood of seeing the evidence if the complementary hypothesis were true [false alarm: $p(e/\bar{h})$], and others. The identification of these concepts offered few problems.

One explanation of the typical response on probabilistic inference word problems of this type is that the subject is ignoring the base rate. To see whether the verbal protocol data are consistent with this explanation, we counted the number of sentences in which the subject mentioned the base rate, after all the information had been presented. This was on the average 1.4 sentences (13% of sentences) for the Cab problem, 4.7 sentences (33%) for the Doctor problem, and 2.3 sentences (14%) for the Twins problem.

To determine whether there is a relation between the mentioning of base rate and its use in producing the answer, the correlations of the number and the proportion of sentences mentioning base rate, with the absolute deviation of the subject's answer from the base rate, was calculated. Findings were: the more sentences that mentioned the base rate, the lower the answer ($r = -.73$, $p = .000$) and the closer the answer to the base rate ($r = -.61$, $p = .001$); the more sentences the subject said, the lower the answer ($r = -.60$, $p = .001$) and the closer the answer to the base rate ($r = -.55$, $p = .003$); the higher the proportion of sentences mentioning the base rate, the lower the deviation of the answer from the base rate (-.13, $p = .280$) and the lower the answer (-.22, $p = .156$).

In sum, the more the subject talked (and presumably, thought) about the problem, the lower the answer (and the closer to the base rate). The effect of the mentioning of the base rate, per se, is secondary, though talking about the base rate seems to bring the answers closer to the base rate. This analysis suggests that the joint use of information processing (protocol) analysis and input/output analysis may be fruitful (see Einhorn, Kleinmuntz, and Kleinmuntz, 1979).

A second question is whether the subjects consider the possibility that the hypothesis might be false (e.g., the Green cab might be responsible for the accident), and further, the possibility that the evidence (cab called "Blue" by witness) might occur if the cab were Green. Only one subject (of ten) mentioned this idea, on one problem. This points to a blind spot in naive subjects' considerations on probabilistic inference word problems (see Doherty, Mynatt, Tweney, and Schiavo, 1979). This is an opportunity for training, and a possible point of contrast between novice and expert behavior.

3.3.2. The main protocol analysis study.

The study, "Interpretation of conditional probabilities in probabilistic inference word problems" (Hamm and Miller, 1988), continues the comparison among the three theories that were tested in (Hamm, 1987b) -- purposeful neglect of base rate information, interpolation between base rate and the 100% which denotes complete acceptance of the evidence, and confusion between conditional expressions of reliability, p(e/h) and p(h/e). Four methods were used here -- manipulation of stimulus information, process tracing, protocol analysis, and recall analysis. The methods are described in "Coder's materials and reliabilities for analysis of thinking aloud protocols from study on the use of conditional probabilities in probabilistic inference" (Hamm, Lusk, Miller, Smith, and Young, 1988). In addition, subjects' behavior was compared not only with the quantitative standard, Bayes' Theorem, but also with six principles expressing qualitative relations between the given information and the answers. Bayes' Theorem has these same qualitative relations.

The qualitative relations consistent with the Bayes' Theorem standard. If people do not both know Bayes' Theorem and have computational tools, then exact application of the formula can not be expected (von Winterfeldt and Edwards, 1986). Nonetheless, we can inquire whether their behavior has qualitative features that are consistent with the prescriptions of Bayes' Theorem. These are, at minimum, that the impacts of each kind of relevant information be in the right direction. Specifically, for the two-hypothesis case,

1. If there is evidence, then it ought to make the subject consider the hypothesis it supports to be more likely.

2. If there are good *a priori* reasons to believe a hypothesis, then the higher this prior probability, the more likely the subject should consider that hypothesis to be; as a special case, if there is relative frequency information about what usually happens, then the higher this base rate, the more likely the subject should consider the hypothesis to be.

3. If in addition to evidence, there is information about the reliability of the evidence, such as information about how frequently the particular evidence would be seen if the hypothesis that the evidence points to were true, then the higher this conditional probability, the more likely the subject should consider the hypothesis to be.

4. If in addition to evidence, there is a second type of information about the reliability of the evidence, which is how frequently the particular evidence would be seen if the complementary hypothesis (the one that the evidence seems to contradict) were true, then the higher this conditional probability, the less likely the subject should consider the hypothesis to be.

5. If evidence is accompanied by both prior probability (or base rate) information and information about the relation between the hypothesis and the evidence, then all this information should be integrated.

6. It is possible to make reasonable assumptions about information that has not been

specified.

These qualitative relations are also required by the alternatives to Bayes' Theorem, with the possible exception of Cohen's "inductive probabilities" (Cohen, 1977; see Schum, 1987).

In this study, it was possible to select among the theories that explain people's performance on the diagnostic class of probabilistic inference word problems. The evidence favors the Confusion hypothesis, which holds that the conditional probability expressing the reliability of the evidence [p(e/h), the probability of observing a particular piece of evidence given that a particular hypothesis is true], is used as the response because people confuse it with the conditional probability p(h/e) [the probability that the hypothesis is true given the evidence has been observed]. The problem asks for the latter, but people think the former is an appropriate answer (Eddy, 1982; Dawes, 1986). This explains why people seem to neglect the base rate information in these problems, even though they may attend to it in other situations. It also gives a basis for predicting when people will follow the qualitative Bayesian principles. Analysis of the conditions of production and utilization of reliability information, in conjunction with what we know about the qualitative Bayesian principles, provides a framework for attempts to improve performance through decision aiding and training.

Influence of formal training on people's vulnerability to conditional probability confusion. Probability is an abstract symbolic language. Although it has entered everyday discourse in a number of realms (e.g., sports, weather), the interpretation of the conditional probabilities presented in these word problems may require special education. We compared undergraduates, presumably untrained in probabilistic inference, with mathematics graduate students, who have studied the formal Bayes' Theorem technique in their probability courses. Only 4 of the 14 mathematics graduate students mentioned the applicability of Bayes' Theorem to the problems. None of them applied it perfectly. Three of them multiplied p(h) by p(e/h), but neglected the possibility of a false alarm, p(e/¬h)*p(¬h), and so missed the point of the ratio in Bayes' Theorem. Thus despite their training they violated the 4th qualitative Bayesian principle (to attend to false positive), although they obeyed the 3rd (to attend to reliability). The fourth subject who attempted Bayes' Theorem mistakenly interpreted the presented p(h/e) conditional probability as p(e/h) in an otherwise correct application of the formula.

Besides their attempts to apply Bayes' Theorem, the mathematics graduate students had statistically more sophisticated intuitions about the problems, as would be expected (see Nisbett, Krantz, Jepson, and Kunda, 1983). They tended to ask for the base rate first [Principle 2] when required to select the order in which to receive the information in the process tracing technique, while undergraduates asked for evidence first [Principle 1]. They tended to adopt the cognitively complex strategy of using base rate in conjunction with other information, while undergraduates either used the base rate alone or ignored it completely. Thus the graduate students were more likely to obey the 5th qualitative Bayesian principle: that evidence and base rate should be integrated.

Implications. We conclude our summary of Study III by considering the implications of our findings concerning whether people behave in accord with the qualititative Bayesian principles. We have shown that people find it natural to attend to the given evidence (Principle 1) on these sorts of word problems, despite the prejudice and stereotypy that appear in other contexts. We have confirmed the findings (Ofir, 1988, and others) that people appreciate the pertinence of base rate information (Principle 2) in most cases. This contradicts the argument of Cohen (1981). We have shown that the apparent exception, the neglect of base rate (Bar-Hillel, 1980; Tversky and Kahneman, 1982), is due to people's difficulties in interpreting conditional probability expressions of reliability, rather than to a lack of appreciation of the base rate. Subjects recognized the pertinence of the reliability of the evidence (Principle 3), although they did not know how to use probabilistic measures of reliability. They spoke about the possibility of a false alarm p(e/¬h) more frequently than about p(e/h) when both ideas were presented to them, which shows that they recognize the pertinence of Principle 4. However, 3 of 4 graduate students who tried to use Bayes' Theorem left out the part that deals with false alarms, and previous work (Doherty, Mynatt, Tweney, and Schiavo, 1979) suggests they don't consider false alarms unless prompted. Subjects often did not integrate information about the base rate and unreliable evidence (Principle 5). They occasionally used numbers that represent attending to the base rate alone, or to the evidence alone. Even when their responses were in between the base rate and 100%, this often represented the use of a conditional probability p(e/h) that they misrecognized as p(h/e), rather than a subjective integration of the competing types of information. Finally, Principle 6 holds that people should make reasonable assumptions concerning information that is missing. Although we did not address this directly in the study, subjects seemed to make simplifying assumptions, rather than making their problems more complex by considering factors that have not been mentioned and assigning reasonable values to them. For example, on only a third of the problems did the subjects spontaneously consider the reliability of the evidence, before the specific reliability information was given to them.

In summary, people tend to follow the qualitative principles at the top of the list better than they followed the later ones. Their failures may be traced to (a) not knowing the principle or not knowing a procedure for carrying it out, (b) not being reminded of the principle, or (c) misinterpreting information (conditional probabilities) and deciding that a principle (Principle 2) is no longer applicable given that misinterpretation. This view has implications for improving performance through training and decision aiding, and for evaluating the performance of any human or man-machine inferencing system.

Training. It should be possible to conduct training in probabilistic inference by building on the basic appreciation that people have for evidence, base rate, and reliability information. Emphasis should be placed on three areas:

1. People should be made alert to the possibility of false positive evidence, which they often neglect.

2. People should be taught how to integrate prior expectations and current evidence, either through Bayes' Theorem or through appropriate estimation techniques that are responsive to the reliability of the evidence.

3. People should be taught to correctly interpret and use reliability information, including avoiding the errors of

    a. misconceiving a p(e/h) probability as a p(h/e), and

    b. assuming that a p(h/e) produced elsewhere is applicable to the present situation.

   In this way they will not be induced to ignore base rate information.

Aids such as the 2 by 2 table explored by Lichtenstein and MacGregor (1984) can sharpen the distinction between the conditional probabilities and also remind people of the possibility of false positive evidence.

Development of expertise. The long term goal of fostering expertise can be distinguished from the short term goal of training someone to perform a particular decision making task. Dreyfus and Dreyfus (1986; see Hamm, 1988b) describe how expertise is developed through repeated engagements with a task in which an analytic perspective is provided by teachers (e.g., Bayes' Theorem in the case of probabilistic inference). They warn against the illusion that intuitive expert performance can be accurate without this sort of long term analytical engagement.

Decision aiding. Decision aids should be designed so that they do not give users the opportunity to misinterpret conditional probabilities and thus enter wrong information into the system. Otherwise. the basic decision aiding approach seems well founded: to remind people about prior expectations. the possibility of false alarms. and the unreliability of evidence (or to automatically take these into account), and to reliably apply the probability calculus if the needed information is available. There is a need for aids in situations where complete information about the probabilistic structure of the environment is not available. The qualitative Bayesian principles may be useful for building such aids. However, decision aids also need to be sensitive to variations in the universality of p(e/h) and p(h) statistics. In addition, non-Bayesian techniques such as Collins and Michalski's (1987) plausible inference or Fox, O'Neil, Glowinski, and Clark's (1988) symbolic inference schemata are being developed in an attempt to provide alternative ways to make inferences using data bases that lack well-measured relative frequencies and conditional probabilities.

Evaluation. It is increasingly becoming necessary to evaluate the inference processes of individuals, groups, or man-machine systems. The results of this study suggest that it might be useful not only to ask whether inference is consistent with Bayes' Theorem or at least with the qualitative Bayesian principles, but also whether appropriate distinctions between the different classes of conditional probability are being made.

4. Follow-up Study using insurance problems.

To examine the effects of substantive expertise and of formal expertise beyond that of mathematics graduate students, and of the combination, a study is being undertaken (with Richard S. Ling) using three problems in the field of insurance. These problems present more details than the problems in the other studies. Considerable effort has been made to assure that the probability facts are accurate and that the decisions and procedures are realistic for the insurance industry context.

The problems are being presented to experienced insurance agents (for their expertise in the content of the problems), to university faculty members in the decision sciences (for their probability expertise), and to actuaries (who have both substantive and probability expertise). Two methods are used in this questionnaire study. To trace the subjects' information seeking strategies, they are required to rate the usefulness of a number of possible types of information, before they receive each additional fact. In addition, the context of the information is varied between subjects: some are given p(e/h) and others are given p(h/e).

This will allow us to discover whether the more experienced and formally better trained subjects value the base rate information and confuse the conditional probabilities in the same way that the undergraduates and mathematics graduate students in the earlier studies did.

References.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. Acta Psychologica, 44, 211-233.

Beyth-Marom, R., and Fischhoff, B. (1983). Diagnosticity and pseudodiagnosticity. J. of Personality and Social Psychology, 45, 1185-1195.

Birnbaum. M.H., and Mellers, B.A. (1983). Bayesian inference: Combining base rates with opinions of sources who vary in credibility. J. Personality and Social Psychology, 45, 792-804.

Cohen. L.J. (1977). The Probable and the Provable. Oxford: Oxford University Press.

Cohen. L.J. (1981). Can human irrationality be experimentally demonstrated? (with open peer commentary) Behavioral and Brain Sciences. 4. 317-370.

Cohen, M., Schum, D.A., Freeling, A.N.S., and Chinnis, J.O., Jr. (1985). On the art and science of hedging a conclusion: Alternative theories of uncertainty in intelligence analysis (Technical Report 84-6). Reston, VA: Decision Science Consortium.

Collins. A., and Michalski, R. (1987). The logic of plausible reasoning: A core theory. Cambridge. MA: Bolt. Baranek, and Newman, Inc.

Dawes. R.M. (1986). Representative thinking in clinical judgment. Clinical Psychology Review. 6. 425-441.

Doherty. M.E., Mynatt. C.R., Tweney, R.D., and Schiavo, M.D. (1979). Pseudodiagnosticity. Acta Psychologica, 43, 111-121.

Dreyfus. H.L., and Dreyfus, S.E. (1986). Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. New York: The Free Press.

Eddy. D.M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases (pp. 249-267). Cambridge: Cambridge University Press.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), Formal representation of human judgment. New York: Wiley, pp 17-52.

Einhorn, H.J., Kleinmuntz, B., and Kleinmuntz, D.N. (1979). Linear regression and process-tracing models of judgment. Psychological Review, 86, 465-485.

Fischhoff. B., and Bar-Hillel. M. (1984). Focusing techniques: A shortcut to improving probability judgments? Organizational Behavior and Human Performance. 34, 175-194.

Fischhoff, B., Slovic, P., and Lichtenstein, S. (1979). Subjective sensitivity analysis. Organizational Behavior and Human Performance, 23, 339-359.

Fogarty. W.M. (1988). Formal investigation into the circumstances surrounding the downing of a commercial airliner by the USS Vincennes on 3 July 1988. Washington, DC: Department of Defense. July 28, 1988.

Fox, J., O'Neil, M., Glowinski, A.J., and Clark, D. (1988). A logic of decision making. Illinois Interdisciplinary Workshop on Decision Making, Champaign-Urbana, Illinois, June.

Hamm, R.M. (1987a). Diagnostic inference: People's use of information in incomplete Bayesian word problems. (Publication #87-11.) Institute of Cognitive Science, University of Colorado, Boulder.

Hamm, R.M. (1987b). Explanations of the use of reliability information as the response in probabilistic inference word problems. (Publication #87-13.) Institute of Cognitive Science, University of Colorado, Boulder.

Hamm, R.M. (1987c). A model of answer choice on probabilistic inference word problems. Society of Mathematical Psychology meetings, Berkeley, CA.

Hamm, R.M. (1988a). Accuracy of probabilistic inference using verbal *vs* numerical probabilities. Psychonomics Society meetings, Chicago.

Hamm, R.M. (1988b). Clinical intuition and clinical analysis: Expertise and the cognitive continuum. In Jack Dowie and Arthur Elstein (Eds.), Professional Judgment: A Reader in Clinical Decision Making. Cambridge: Cambridge University Press, pp 78-105.

Hamm, R.M. (1988c). Evaluation of a method of verbally expressing degree of belief by selecting phrases from a list. (Publication #88-6.) Institute of Cognitive Science, University of Colorado, Boulder.

Hamm, R.M., Lusk, C.M., Miller, M.A., Smith, D.L., and Young, I.E. (1988). Coder's materials and reliabilities for analysis of thinking aloud protocols from study on the use of conditional probabilities in probabilistic inference. (Publication #88-13.) Institute of Cognitive Science, University of Colorado, Boulder.

Hamm, R.M., and Miller, M.A. (1988). Interpretation of conditional probabilities in probabilistic inference word problems. (Publication #88-15.) Institute of Cognitive Science, University of Colorado, Boulder.

Hammond, K. R. (1981). Principles of organization in intuitive and analytical cognition. (Report #231). Boulder: University of Colorado, Center for Research on Judgment and Policy.

Kahneman, D., and Tversky, A. (1972). On prediction and judgment. Oregon Research Institute Research Monograph, 12(4).

Lichtenstein, S., and MacGregor, D. (1984). Structuring as an aid to performance in base rate problems. Technical Report 84-16. Eugene, OR: Decision Research.

Lyon, D., and Slovic, P. (1976). Dominance of accuracy information and neglect of base-rates in probability estimation. Acta Psychologica, 40, 287-298.

MacGregor, D., Lichtenstein, S., and Slovic, P. (1988). Structuring knowledge retrieval: An analysis of decomposed quantitative judgments. Organizational Behavior and Human Decision Processes, 42, 303-323.

Niiniluoto, I. (1981). L.J. Cohen versus Bayesianism. Behavioral and Brain Sciences, 4, 349.

Nisbett, R.E., Krantz, D.H., Jepson, C., and Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. Psychological Review, 90, 339-363.

Ofir, C. (1988). Pseudodiagnosticity in judgment under uncertainty. Organizational Behavior and Human Decision Performance, 42, 343-363.

Ofir, C., and Lynch, J.G., Jr. (1984). Context effects on judgment under uncertainty. J. Consumer Research, 11, 668-679.

Robinson, Lori B., and Hastie, Reid. (1985). Revision of beliefs when a hypothesis is eliminated from consideration. Journal of Experimental Psychology: Human Perception and Performance, 11, 443-456.

Schum, D.A. (1987). Evidence and Inference for the Intelligence Analyst. Lanham, MD: University Press of America.

Schum, D.A., and Martin, A.W. (1980). Probabilistic opinion revision on the basis of evidence at trial: A Baconian or a Pascalian process? (Report No. 80-02). Houston, Texas: Rice University.

Shafer, G. (1976). A Mathematical Theory of Evidence. Princeton, NJ: Princeton University Press.

Shortliffe, E. (1984). Coming to terms with the computer. In S. J. Reiser and M. Anbar (Eds.), The machine at the bedside: Strategies for using technology in patient care. Cambridge: Cambridge University Press, pp 235-240.

Slovic, P., and Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 6, 649-744.

Tversky, A., and Kahneman, D. (1982). Evidential impact on base-rates. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), Judgment under Uncertainty: Heuristics and Biases. New York: Cambridge University Press, pp. 153-160.

von Winterfeldt, D., and Edwards, W. (1986). Decision analysis and behavioral research. New York: Cambridge University Press.

~~Contract MDS-903-86-K-0265~~

~~September 1, 1986 to February 28, 1989~~

PEOPLE MISINTERPRET CONDITIONAL PROBABILITIES:

~~FINAL REPORT OF PROJECT~~

~~USING PROTOCOL ANALYSIS AND PROCESS TRACING TECHNIQUES,~~

~~TO INVESTIGATE PROBABILISTIC INFERENCE.~~

Robert M. Hamm

~~Institute of Cognitive Science~~

~~Campus Box 345~~

University of Colorado

~~Boulder CO 80309-0345~~

~~Robert M. Hamm, Project Director~~

~~(303) 492-2936~~

~~Institute of Cognitive Science~~

~~Publication Number 89-4~~

Army Research Institute for the
Behavioral and Social Sciences

~~Final Report.~~

~~March 20, 1989~~