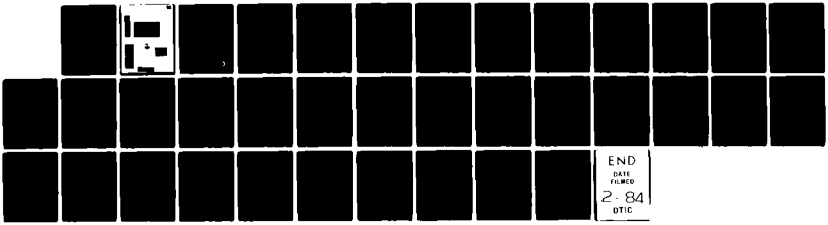
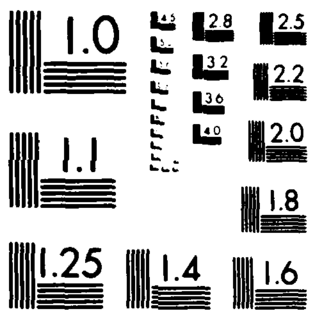
AD-A137 032 A THEORY OF DIAGNOSTIC INFERENCE: CONTRACT FINAL REPORT 1/1
(U) CHICAGO UNIV IL CENTER FOR DECISION RESEARCH
H J EINHORN ET AL. NOV 83 TR-8 N00014-81-K-0314

UNCLASSIFIED F/G 5/10 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

D

(12)

A THEORY OF DIAGNOSTIC INFERENCE: CONTRACT FINAL REPORT

Hillel J. Einhorn and Robin M. Hogarth
University of Chicago
Graduate School of Business
Center for Decision Research

November 1983

Sponsored by:
Office of Naval Research
Contract Number, N00014-81-K-0314
Work Unit Number, NR 197-071

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 8	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Theory of Diagnostic Inference: Contract Final Report		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Hillel J. Einhorn and Robin M. Hogarth		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0314
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Decision Research, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, Illinois 60637		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 197-071
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217		12. REPORT DATE November 1983 13. NUMBER OF PAGES 35
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

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)

Ambiguity, causal judgment, decision making, diagnosis, probabilistic inference, and uncertainty.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report summarizes research aimed at developing a theory of diagnostic inference carried out during the 30-month contract period of April 1, 1981 through September 30, 1983. The essential aspects of diagnostic inference are that they are causal rather than correlational, — pauce

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

con't

20.

backward rather than forward (one goes from effects to prior causes), concerned with a specific rather than the general case, and constructive (one can synthesize, enlarge, or otherwise develop new hypotheses). We have developed and tested models of two aspects of diagnostic infer- inference: (1) judgments of the causal strength of a hypothesis/ explanation, and (2) how probabilistic assessments concerning the occurrence of a past event are made on the basis of (often conflicting) evidence received from less than perfectly reliable sources. The psychological rationales, quantitative formulations, and implications of these models are presented in this report together with a description of various experiments designed to test the models. We also discuss several commonalities between the two lines of research: the use of cognitive anchoring-and-adjustment strategies to cope with complex inference tasks; the constructive nature of diagnosis; the importance of surprise in inference; and the relation of the present work to normative standards of judgment and choice.

TABLE OF CONTENTS

Introduction

This report summarizes our work on developing a theory of diagnostic inference for the period April 1, 1981 through September 30, 1983. By diagnosis, we mean the following: Given the occurrence of a set of outcomes/results/symptoms, people infer what causal process could have produced the observed effects. The essential aspects of such inferences are that they are causal rather than correlational, backward rather than forward (one goes from effects to prior causes), concerned with a specific rather than the general case, and constructive (one can synthesize, enlarge, or otherwise develop new hypotheses). The importance of diagnosis goes beyond its obvious role in making sense of experience; it is crucial for predictive judgment as well as for defining what variables are "relevant." Moreover, since the evidence typically available for making diagnoses is fallible and/or conflicting, the process takes place under uncertainty.

The development of our theory has followed two complementary paths: (1) The formulation and testing of a theory of how causal judgments are made; and (2) the creation of a theory of evidence that concerns how judgments are made in ambiguous situations. We consider each of these topics in turn. Subsequently, we discuss some commonalities between the two lines of research.

Judgments of Causality

We have developed, and experimentally tested, a model of how people judge the causal strength of a hypothesis/explanation. As a specific illustration, imagine that a watch face has been struck sharply by a hammer and the glass breaks. You are then asked to assess how likely the breakage was caused by the force of the hammer. We argue that answers to this question will be mediated by three types of information: (1) The number and strength of specific alternative explanations. Part of the reason that the force of the hammer is a strong causal candidate is due to the fact that it is difficult to imagine specific alternatives that could reduce one's belief in that explanation; (2) The assumed causal background against which the judgment is made. For example, reconsider your response to the above question if the context was changed to a watch factory where a hammer strikes watch faces as part of a testing procedure. In this context, it is more likely that a defect in the glass will be judged to be the cause; (3) The judged causal strength of the explanation. We maintain that people use certain cues-to-causality in assessing the plausibility of an explanation; namely, temporal order, contiguity, covariation, and similarity of cause and effect. In our example, note that the glass broke immediately after being struck by the hammer; there is a high correlation between the breaking (or not) of glass with the force of solid objects; and there is similarity between the length and strength of cause and effect.

The role of alternatives

Causal judgments are complex. We therefore propose that people handle this task sequentially by a cognitive anchoring-and-adjustment strategy. This can be illustrated by considering how people adjust causal beliefs by the number and plausibility of specific alternative explanations. Consider an outcome Y, an initial explanation X, and alternative explanation Z_1. Furthermore, denote the "gross strength" of an explanation as being its plausibility or strength before competing alternatives are considered. Thus, the gross strengths of X and Z_1 refer to their plausibility when each is considered the sole explanation of Y. We propose that people anchor on the gross strength of the initial explanation X, and then adjust downward for the gross strength of Z_1. Moreover, the amount of the adjustment will depend on the strength of the anchor as well as the strength of the alternative. In particular, we assume that alternatives of equal strength discount strong explanations more than weaker ones. For example, imagine that one anchors on a weak hypothesis and is then confronted with a strong alternative. Since the anchor is already low, the size of the adjustment cannot be too large (indeed, if the anchor were worthless, there would be no adjustment). On the other hand, if the anchor was strong, we argue that the same alternative would discount the anchor substantially. Therefore, the basic idea is that the stronger the anchor, the larger the adjustment (holding the strength of alternatives equal). We call the strength of an explanation after it is reduced by an alternative, its "net strength."

The above process can be formally represented as follows:

$$S_1(Y,X|B) = s_o(Y,X|B) - w_o\, s(Y,Z_1|B) \qquad (1)$$

where,

$S_1(Y,X|B)$ = net strength of the causal link of Y with X, conditional on background B, after adjusting for Z_1

$s_o(Y,X|B)$ = gross strength of the causal link of Y with X, conditional on background B

$s(Y,Z_1|B)$ = gross strength of the causal link of Y with Z_1, conditional on background B

w_o = adjustment weight applied to the gross strength of Z_1 $(0 < w < 1)$

In equation (1) we adopt the convention that capital "S" stands for net strength and small "s" denotes gross strength. Of course, before any alternative is considered, $S_o = s_o$. Note that the adjustment weight, w, has the same subscript as the anchor since it is a function of the latter. Now consider what happens when a second alternative, Z_2, is introduced. We assume that the anchor-and-adjust strategy proceeds sequentially so that the net strength of X becomes the new anchor for the next adjustment. Thus,

$$S_2(Y,X|B) = S_1(Y,X|B) - w_1\ s(Y,Z_2|B) \tag{2}$$

Equation (2) can be be generalized to account for the net strength of X after the kth alternative (k = 1,2, ..., K); thus,

$$S_k(Y,X|B) = S_{k-1}(Y,X|B) - w_{k-1}\ s(Y,Z_k|B) \tag{3}$$

Furthermore, since $S_k(Y,X|B)$ is a judged likelihood, it is bounded between 0 and 1.

Now consider the functional relation between the strength of the anchor and the adjustment weight, w (called the "adjustment weight function"). It was assumed above that stronger anchors have larger adjustments. This implies that the adjustment weight is a monotonically increasing function of the strength of the anchor. To see this, consider equation (3) when the gross strength of Z_k is constant and the anchor varies in strength. It is clear that as $S_{k-1}(Y,X|B)$ increases, w_{k-1} must also increase to give larger adjustments. To model this monotonic relation, we posit a simple and convenient form, although others might serve as well; thus,

$$w_{k-1} = [S_{k-1}(Y,X|B)]^{\alpha} \quad (\alpha > 0) \tag{4}$$

Equation (4) is illustrated in Figure 1. Note that α affects the amount by

Insert Figure 1 about here

which explanations are discounted and can thus be thought of as varying as a function of task and/or individual characteristics. For example, $\alpha > 1$ implies that the adjustment weights are less than the anchor and corresponds to underweighting the impact of disconfirming evidence; $\alpha = 1$ implies that adjustment weights equal the anchor; $0 \leq \alpha < 1$ implies that adjustment weights are larger than the anchor. The importance of this for the final net strength of X can be seen by first substituting (4) into (3). This yields:

$$S_k(Y,X|B) = S_{k-1}(Y,X|B) - [S_{k-1}(Y,X|B)]^{\alpha} \, s(Y,Z_k|B) \qquad (5)$$

which can be thought of as the computational form of the anchor-and-adjust model. Indeed, as we have demonstrated empirically (see below), when α is estimated from data, equation (5) can be used to predict how people revise their causal judgments.

To illustrate how the model specified in (3) and (5) captures important aspects of the causal judgment process, consider the model in its non-sequential form;

$$S_K(Y,X|B) = s_o(Y,X|B) - \sum_{k=1}^{K} w_{k-1} \, s(Y,Z_k|B) \qquad (6)$$

That is, the net strength of an explanation is equal to its gross strength minus the sum of the adjusted alternative explanations. In other words, we posit that net strength follows a difference model as opposed to a ratio model such as probability theory. This means that net strength can be low when there are no alternatives if the gross strength of X is itself low.

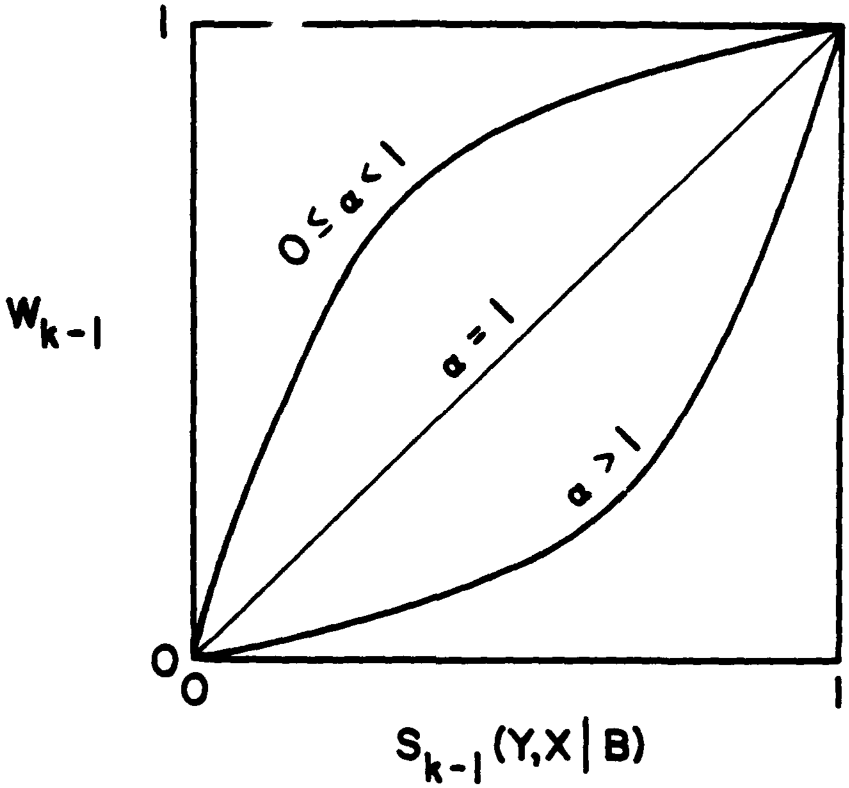


Figure 1. The Adjustment Weight Function

Moreover, net strength can also be low when gross strength is high if there are many strong alternatives. Indeed, net strength can only be high if gross strength is high and the strength of specific alternatives is low. For example, reconsider the initial watch-hammer scenario and contrast the net strength of the "force of the hammer" explanation with the net strength of any single explanation for the following questions:

1. Why are the outer rings of Saturn braided?
2. Why was Ronald Reagan elected President in 1980?

For the first question, it is difficult to generate a single explanation, thus suggesting its gross strength is low However, although there are no competing explanations, net strength remains low in accord with equation (6). For the second question, there are many strong explanations (e.g., the situation of the economy; the rise of the moral majority; the unresolved Iranian hostage problem; etc.). Therefore, while the gross strength of these are high, the net strength for any single one is low precisely because the others are plausible alternatives. On the other hand, the watch-hammer question leads to high net strength since the explanation is strong and there are few plausible alternatives. In short, it is argued that like good patterns, good explanations have few alternatives; or, to be more precise, whereas good explanations imply few alternatives, the lack of alternatives does not imply good explanations.

The causal background

Above, we have been careful to condition all terms on the causal background, B. We do this because diagnostic inference is typically invoked to make sense of deviations via causal explanation. However, the meaning of a deviation is itself crucially dependent on some assumed background or field. Specifically, we argue that causal relevance is generally related to the degree that a variable is a difference-in-a-background. By this is meant that factors that are part of some presumed background are judged to be of little or no causal relevance. For example, does birth cause death? While the former is both necessary and sufficient for the latter (and thus covaries perfectly with it), it seems odd to consider one the cause of the other. The reason is that death presumes that one has been born. Therefore, "birth" is part of the background and its causal relevance is low.

Cues-to-causality

In our model, the judged plausibility of a hypothesis/alternative (i.e., gross strength) is based on cues-to-causality. These are postulated to be imperfect indicators of causal relations that combine and trade-off in judgments of gross strength. Specifically, we define gross strength to be the following function of the cues-to-causality;

$$s(Y,X|B) = Q_1\gamma(\lambda_2 Q_2 + \lambda_3 Q_3 + \lambda_4 Q_4) \tag{7}$$

where,

Q_1 = temporal order = (0,1)

Q_2 = contiguity

Q_3 = covariation

Q_4 = similarity

$$\gamma = \begin{cases} 0 \text{ if } Q_4 < \text{threshold} \\ 1 \text{ if otherwise} \end{cases}$$

λ_i = importance weight for the ith cue (i = 1, ..., 4)

Note that if either temporal order is inappropriate or similarity is below threshold, gross strength is zero. Otherwise, the cues of contiguity, covariation, and similarity will trade-off.

Experimental evidence

Three types of experiment have been performed to test (1) the role of alternatives via the sequential anchor-and-adjust model, (2) the use of the cues-to-causality, and (3) the causal field concept.

(1) Equation (5) was estimated and tested on data where subjects were first asked to judge the gross strength of a hypothesis, and then assess its net strength after sequential presentation of two specific alternatives. In fact, the initial gross strength judgment and first assessment of net strength were used to estimate the α parameter, and the second net strength judgment was predicted on this basis. Across different permutations of hypotheses and alternatives in two scenarios, the mean absolute deviation was a mere .02 for judgments made on a 0 to 1 scale.

(2) A second series of experiments tested the role of the cues-to-causality as operationalized in equation (7). Three cues, contiguity, covariation, and similarity were varied factorially across 8 scenarios. Results showed predicted main effects for covariation and similarity but not contiguity. Furthermore, the data revealed interesting interactions between the cues and the scenarios thereby emphasizing the notion that the cues are perceived conditionally on the context or causal field in which they are embedded. In a second experiment, when the contiguity cue was manipulated to be more salient in the scenarios, significant main effects of the predicted sign were observed. In these experiments subjects not only judged a causal candidate, but made subsequent judgments after being informed of the presence of specific alternatives. This permitted two further tests of our model: (a) predictions using equation (5); and (b) a test of the similarity threshold hypothesis implicit in equation (7). In both cases, our hypotheses were supported by the data.

(3) Recall the example given above of the hammer hitting the face of a watch when no explicit context is provided and when the scene is supposed to take place in a watch factory. The effect of this shift in the causal background, with its corresponding change in the strength of causal candidates (the force of the hammer vs. a defect in the glass) was tested experimentally in both a between- and within-subjects design. The results provided strong support for the notion that shifts in the background can have dramatic effects on attributions of causality.

To summarize, the three components of our model (role of alternatives, cues-to-causality, and background) were tested and found to support our conceptualization. Whereas we clearly do not claim that our model provides a complete picture of the myriad issues involved in causal judgment, it does provide a solid, and parsimonious foundation on which to build. Furthermore, our model can be shown to subsume, and even quantify, earlier attempts to conceptualize causal judgment, e.g., various attribution theories

Evaluating Evidence under Ambiguity

An important input to diagnostic inference is the evaluation of evidence. Moreover, this usually takes place under conditions of uncertainty thereby invoking probabilistic reasoning. However, in an important paper Daniel Ellsberg demonstrated that subjective probabilities inferred from choices amongst gambles do not necessarily conform to the axioms of probability theory. Specifically, Ellsberg showed that such violations are likely to occur in ambiguous conditions where lack of knowledge about the process generating outcomes induces uncertainty about one's own degree of uncertainty. Contrast, for example, the nature of the uncertainty one faces in choosing between heads or tails on the flip of a fair coin as opposed to the uncertainties involved in deciding between two candidates for President.

<u>Task description</u>

The task we have explored requires subjects to assess the likelihood that a particular event happened given f reports favoring its occurrence and c reports favoring an alternative. The reports are deemed to emanate from a single source (e.g., witnesses) and to carry equal weight. A model of the task is depicted in Figure 2. That is, (1) an event occurs, (2) it is sensed by a source that can, in principle, be characterized by levels of sensitivity

Insert Figure 2 about here

and bias, (3) the source decides what to report (this can be thought of as analogous to a signal detection task), (4) several reports are obtained, and (5) the judge combines the reports with expectations based on the content of the scenario to come up with a likelihood judgment, S(f:c). It is important to note that in our task no explicit information is given to the subjects about the reliability (sensitivity, bias) of the reporting source. Rather, *this must be inferred* from the content of the scenarios. Thus, expectations held by the subjects (5) are inferred from ambiguous stimuli.

The importance of ambiguity in our task can be emphasized by contrasting it with more commonly studied probabilistic paradigms: (a) whereas we are dealing with inference from unreliable sources, this is not the standard cascaded inference task in that the judge does not know the precise level of the source's reliability; (b) we cannot apply the theory of signal detection since too many parameters are unknown (e.g., the source's hit-rate and false-alarm rate). We have, instead, a case of several observers reporting on one event (trial) as opposed to multiple trials of the same observer; and (c) the standard Bayesian paradigm cannot be applied *since no values are given either* for the prior probabilities of events or source reliability. However, since precise probabilities are not available in many important real-world problems,

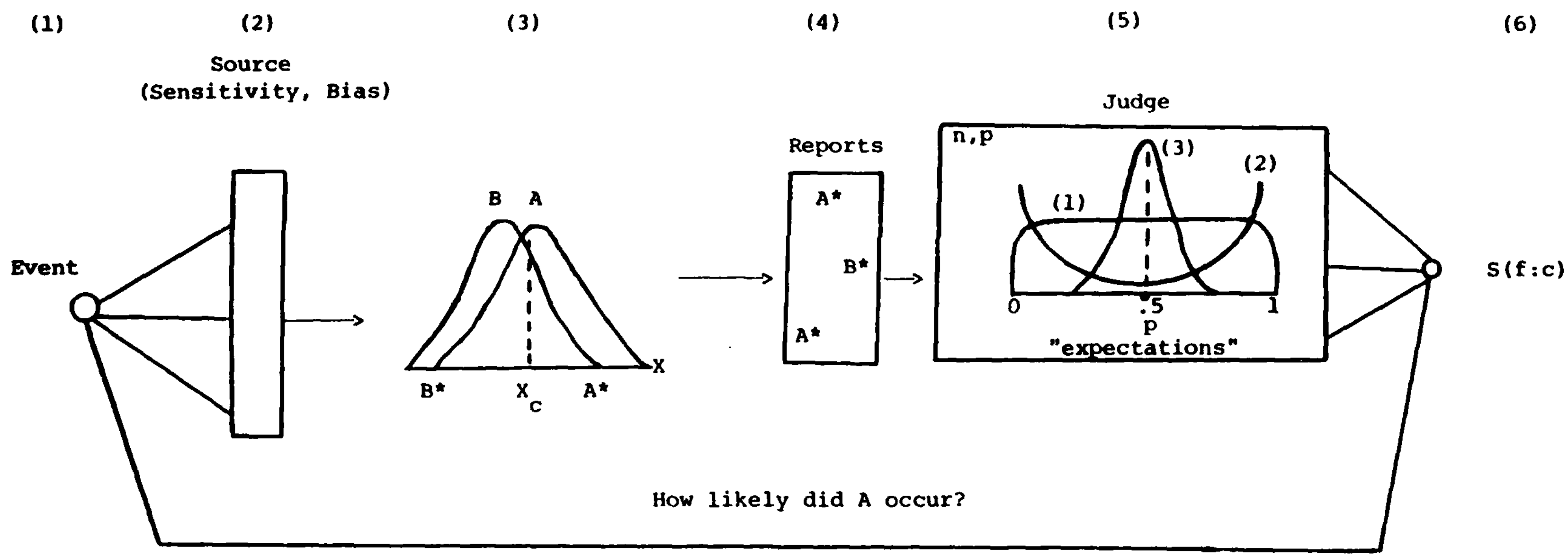


Figure 2. Structure of the ambiguous inference task

we believe that the ambiguity inherent in our task is both realistic and important.

A model for ambiguous inferences

Our model of how people handle the above task involves anchoring on the evidence presented ("what is") and adjusting for outcomes that one could imagine occurring ("what might have been"). That is;

$$S(f:c) = p + k \tag{8}$$

where $p = f/n$, the proportion of reports favoring the hypothesis, and k captures the net effect of imagining (i.e., mentally simulating) alternative outcomes greater and smaller than p. Specifically, we postulate that k will be affected by four variables: (1) the level of p since if $p = 1$, $k < 0$; but if $p = 0$, $k > 0$; (2) n, the number of reports; (3) a parameter θ that reflects the perceived lack of credibility of the source and dissimilarity of the signals $(0 < \theta < n)$; and (4) the effect of differentially weighting imagined outcomes greater and smaller than p. This is captured by a parameter β. The full model is;

$$S(f:c) = p + \frac{\theta}{n}(1 - p - p^{\beta}) \tag{9}$$

and is illustrated in Figure 3. Note that the absolute size of the adjustment

Insert Figure 3 about here

from p increases in θ but decreases in n. Note further that $S(f:c)$ is regressive with respect to p. Moreover, the parameter β implies a unique value of p, denoted p_c, for which $k = 0$ and thus where $S(f:c) = p$. We interpret β as the individual's "attitude toward ambiguity in the circumstances." In summary, note that when $(f/n) = p < p_c$, $k > 0$ so that $S(f:c) > p$; however, when $(f/n) = p > p_c$, $k < 0$ so that $S(f:c) < p$.

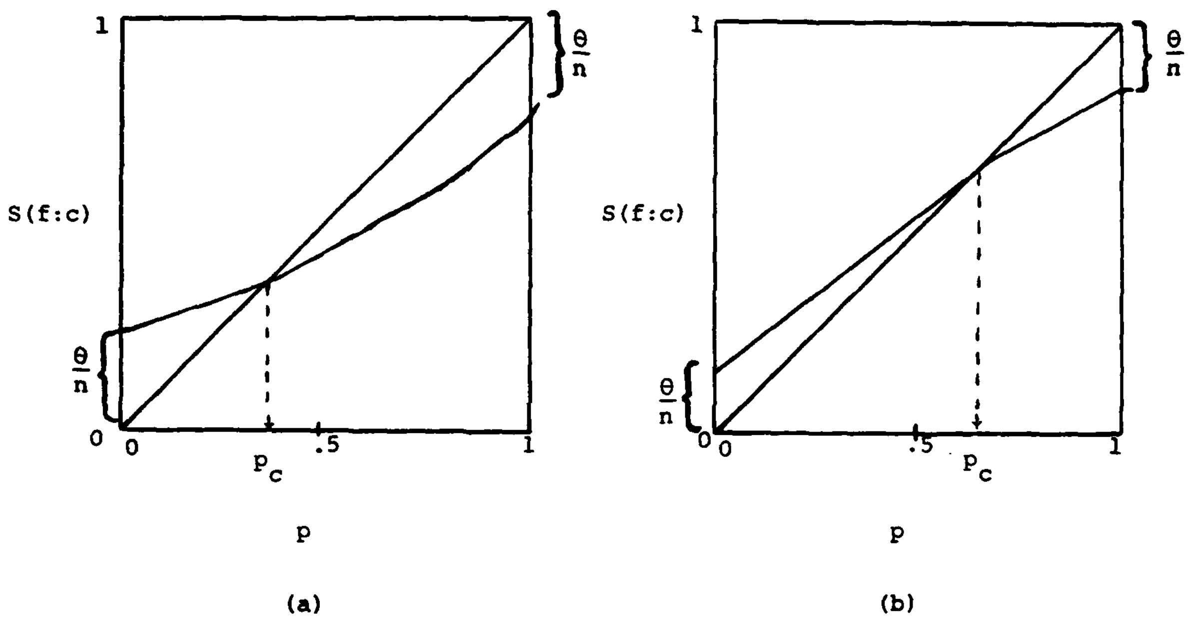


Figure 3. Two examples of S(f:c) as a function of p. In (a), $\beta < 1$ implies $p_c < .5$; in (b), $\beta > 1$ implies $p_c > .5$.

Experimental evidence

We have tested our model and its implications in a series of four experiments.

(1) Using a scenario involving a hit-and-run accident and witnesses' conflicting responses as to the color of the offending car, we have tested and fitted equation (9) on both group and individual data. We have (a) demonstrated how subjects trade-off p and n in assessing the impact of evidence, and (b) as implied by our model, subjects' judgments of the probabilities of complementary events do not necessarily sum to one. More specifically, the parameters of our model (i.e., θ and β) define the conditions under which such judgments will be additive (i.e., sum to one), sub-additive, or super-additive. Moreover, our data bear out these predictions for both groups and individual subjects.

(2) In a replication experiment we demonstrated the validity of the model over four additional scenarios involving different content.

(3) In a factorial experiment, we manipulated both the credibility and dissimilarity of the sources of evidence in order to induce differences in the θ and β parameters. Source credibility was seen to have a significant effect on both parameters (i.e., on the means of the parameters of individual subjects), but no effect was observed for source dissimilarity. Moreover, these results were obtained in both between- and within-subject designs. In addition, we investigated and found consistent individual differences in subjects' judgmental strategies. That is, subjects' individual θ and β parameters were significantly correlated across tasks involving different scenarios. In addition, the extent to which subjects made adjustments for "what might have been," as measured by their θ parameters, was significantly related to the mean absolute deviations between their actual judgments and

model predictions. We interpret this finding as evidence that the ability to execute a judgmental strategy consistently is inversely related to the amount of mental simulation used in generating responses.

(4) We investigated whether individuals' choices between gambles could be predicted from knowledge of their θ and β parameters estimated from a separate inference task. Not only did we find that such choices between gambles could be predicted quite well, but strength of preference between gambles was related to the θ and β parameters in a theoretically consistent manner.

To summarize, our experimental data support our descriptive model of how people handle the ambiguity inherent in a diagnostic inference task via a cognitive anchoring-and-adjustment strategy. We believe that the success of our modeling efforts underscores the importance of incorporating the effects of ambiguity in the perception of uncertainty as well as on risky choice. Important real-world risks are rarely defined explicitly (i.e., unambiguously) and thus it is essential to understand how people treat ambiguous evidence. Second, in trying to analyze our subjects' inference task from a normative perspective, we were struck by the number of assumptions one would need to make to apply an appropriate model, let alone the need for computation such a model would impose. And yet, our subjects were able to provide consistent responses to the experimental stimuli by essentially striking a compromise between "what was" (i.e., the data) and "what might have been" (i.e., the result of the imagined outcomes) via a quite simple cognitive strategy. In addition, whereas the subjects' judgments deviated consistently from some normative prescriptions, it would be difficult to say that they were not acting sensibly.

Some Commonalities

The development of our theory of diagnostic inference has followed two complementary paths. We now discuss some commonalities between these two lines of research.

(1) Anchoring-and-adjustment strategies. Both our models are based on the notion that people use cognitive anchoring-and-adjustment strategies to handle complex inference tasks. In the causal model, it is assumed that people anchor on the gross strength of the hypothesis and then adjust, sequentially, for the strength of specific alternatives. Moreover, we posit that the adjustment weight at each stage is a function of the associated anchor. In the ambiguity model, there is also an adjustment for alternatives that "might have been." Note that in both models the assumed details of the anchoring-and-adjustment processes have been made quite explicit. More generally, we believe that anchoring-and-adjustment strategies are prevalent in many sequential judgment tasks. Thus, the payoff from investigating such strategies carefully in these tasks could also be substantial in other areas of judgment and choice.

(2) Construction in diagnosis. Diagnosis is largely a constructive activity. This is particularly evident in the causal model when one examines the role of the causal background. People respond very quickly to cues in the causal background that allow them to shift the perspective within which a problem is viewed (recall the hammer and watch face example). In the ambiguity model, expectations and imagination clearly play important roles in simulating "what might have been." Neither of our models could be said to provide an account of construction in the diagnostic process. However, since both models incorporate some aspects of construction, they do provide a foundation upon which this crucial cognitive activity can be investigated.

(3) Surprise in inference. Our models can also be thought of as throwing some light on the importance of surprise in diagnostic inference. Surprise is a reaction to discrepancies between expectations and reality. Thus, in the causal model we note that surprise, in the form of deviations-in-a-field, is an important stimulus to diagnostic activity. Also, since the cues-to-causality can induce expectations, when these are violated the resulting surprise will trigger diagnostic activity and attempts to reinterpret experience. In the context of the ambiguity model, we note that data can create surprise (relative to expectations) either because they are "too good" or "not good enough." Indeed, when data are perceived to reach a given level of unreliability, our model predicts that if data are "too good," the surface meaning will suggest the opposite conclusion. To illustrate, consider your reaction if you were assessing whether a particular car had been speeding (i.e., travelling at more than 55 m.p.h.) and 10 eyewitnesses each independently stated that the car had been travelling at 73 m.p.h. Both the complete agreement of the witnesses and the level of precision of their responses are such that you would probably dismiss the data. From equation (9), and assuming that θ (the source unreliability parameter is at its maximum of n), we obtain

$$S(f:c) = 1 - p^{\beta} \qquad (10)$$

which, since $p = (f/n) = 1$ (all witnesses agreed) becomes 0. In other words, when high source unreliability is attributed to witnesses, our model specifically predicts a constructive rejection of the evidence. On the other hand, our model cannot handle situations where data are "not good enough" since the resolution of this type of surprise requires a creative re-organization of the hypotheses being considered. This extension is one of several challenges for future work.

(4) Normative issues. In research on judgment and decision making it is almost mandatory to compare subjects' responses with normative benchmarks. However, this is problematic in the present research. First, there is no normative theory for judgments of causality. Second, it is unclear what the normative model should be in reference to our model of judgment under ambiguity. To be sure, the latter model deviates from Bayesian prescriptions in two respects: (1) as n increases judgments asymptote at p instead of 0 or 1. However, this result can also be obtained in some cascaded inference models that are based on Bayes' theorem; (2) under certain circumstances, judgments of the probabilities of complementary events do not sum to one in our model. Whereas this possibility might be considered a "mistake," we note that there is currently considerable interest in the normative status of non-additive probabilities, the role of the "weight of evidence" in determining probabilistic belief, and so on. Our descriptive model, however, specifies the precise conditions under which additivity/non-additivity will occur, and in doing so captures how people respond to the "weight of evidence" in their judgments.

Given the lack of a normative theory of causality, what contribution can our model of causal judgment make to assessing the quality of diagnostic inference? Whereas we have no definitive answer to this question, we note that our model does highlight two relevant trade-offs. First, our conceptualization of the cues-to-causality posits that these are imperfect, probabilistic indicators of causal relations. Thus, whereas the cues will indicate appropriate causal relations in the environment, they can also indicate invalid relations thereby inducing mistaken beliefs. For a limited organism dealing with a vastly more complex environment, it is certainly appropriate to deal consistently with a limited number of cues-to-causality, and particularly since these can combine in myriad ways. On the other hand, the use of these cues will inevitably imply a trade-off between acquiring some valid causal knowledge, on the one hand, and some superstitions, on the other.

Second, the causal field and the cues-to-causality both play an important role in limiting the number of interpretations people make in inferential tasks, and thus in creating "order-out-of-chaos." Furthermore, the adoption of a particular background and the use of the cues proceed quickly and are often marked by a lack of awareness that a delimiting process has taken place. The benefits to be gained from such automatized processes are large. However, they come at the cost of reducing the probability that people can achieve more creative representations of inferential tasks. In short, the organism may often face a trade-off between stable representations of problems, on the one hand, and the possibility of generating more creative representations of the same stimuli, on the other.

Finally, whereas this report has described the work accomplished to date in the construction of our theory of diagnostic inference, it should be realized that a comprehensive theory is far from complete. However, we do believe that we now have solid foundations on which to examine both the psychology of causal reasoning and the interpretation of fallible, conflicting evidence. It is on these foundations that we are extending our theoretical and empirical investigations.

Technical Reports and Publications

Technical Reports

1. Einhorn, H. J., and Hogarth, R. M. Prediction, diagnosis and causal thinking in forecasting. University of Chicago, Graduate School of Business, Center for Decision Research, September 1981.

2. Einhorn, H. J., and Hogarth, R. M. A theory of diagnostic inference: I. Imagination and the psychophysics of evidence. University of Chicago, Graduate School of Business, Center for Decision Research, June 1982.

3. Einhorn, H. J., and Hogarth, R. M. A theory of diagnostic inference: II. Judging causality. University of Chicago, Graduate School of Business, Center for Decision Research, September 1982.

4. Einhorn, H. J., and Hogarth, R. M. A theory of diagnostic inference: Judging causality University of Chicago, Graduate School of Business, Center for Decision Research, August 1983.

5. Einhorn, H. J., and Hogarth, R. M. Ambiguity and uncertainty in probabilistic inference. University of Chicago, Graduate School of Business, Center for Decision Research, September 1983.

6. Lipe, M. G. A cross-study analysis of covariation judgments. University of Chicago, Graduate School of Business, Center for Decision Research, April 1982.

7. Klayman, J. Learning in a probabilistic environment: A new approach, and some preliminary findings. University of Chicago, Graduate School of Business, Center for Decision Research, May 1983. (Paper presented at the meeting of the Midwestern Psychological Association, Chicago.)

8. Einhorn, H. J., and Hogarth, R. M. A theory of diagnostic inference: Contract final report. University of Chicago, Graduate School of Business, Center for Decision Research, November 1983.

Publications

Einhorn, H. J., & Hogarth, R. M. Prediction, diagnosis, and causal thinking in forecasting. Journal of Forecasting, 1982, 1, 23-36.

Hogarth, R. M. On the surprise and delight of inconsistent responses. In R. M. Hogarth (Ed.), New directions for methodology of social and behavioral science: Question framing and response consistency. San Francisco, CA: Jossey-Bass, 1982.

Hogarth, R. M. Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. Psychological Bulletin, 1981, 90, 197-217.

Acknowledgements and Scientific Personnel

We wish to thank Martin A. Tolcott, Willard S. Vaughn Jr., and J. Randolph Simpson of the Office of Naval Research for their support and encouragement during the contract period. At the University of Chicago, we received most able support and assistance from Charlesetta Nowels, Haim Mano, Ann McGill, Marlys Lipe, Zvi Gilula, Werner Wothke, John Lyons, and Diane Vitiello. Several colleagues also provided valuable comments and criticisms at various stages of the work, notably Joshua Klayman, William Goldstein, and Jay Russo.

OFFICE OF NAVAL RESEARCH

Engineering Psychology Group

TECHNICAL REPORTS DISTRIBUTION LIST

OSD

CAPT Paul R. Chatelier
Office of the Deputy Under Secretary of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D. C. 20301

Dr. Dennis Leedom
Office of the Deputy Under Secretary of Defense (C^3I)
Pentagon
Washington, D. C. 20301

Department of the Navy

Engineering Psychology Group
Office of Naval Research
Code 442 EP
Arlington, VA 22217 (2 cys.)

Aviation & Aerospace Technology Programs
Code 210
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Communication & Computer Technology Programs
Code 240
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Physiology & Neuro Biology Programs
Code 441NB
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Deparment of the Navy

Tactical Development & Evaluation Support Programs
Code 230
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Manpower, Personnel & Training Programs
Code 270
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Mathematics Group
Code 411-MA
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Statistics and Probability Group
Code 411-S&P
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Information Sciences Division
Code 433
Office of Naval Research
800 North Quincy Street
Arlington, VA 2217

CDR K. Hull
Code 230B
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Department of the Navy

Special Assistant for Marine Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Dr. J. Lester
ONR Detachment
495 Summer Street
Boston, MA 02210

Mr. R. Lawson
ONR Detachment
1030 East Green Street
Pasadena, CA 91106

CDR James Offutt, Officer-in-Charge
ONR Detachment
1030 East Green Street
Pasadena, CA 91106

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D. C. 20375

Dr. Michael Melich
Communications Sciences Division
Code 7500
Naval Research Laboratory
Washington, D. C. 20375

Dr. J. S. Lawson
Naval Electronic Systems Command
NELEX-06T
Washington, D. C. 20360

Dr. Robert E. Conley
Office of Chief of Naval Operations
Command and Control
OP-094H
Washington, D. C. 20350

CDR Thomas Berghage
Naval Health Research Center
San Diego, CA 92152

Department of the Navy

Dr. Robert G. Smith
Office of the Chief of Naval Operations, OP987H
Personnel Logistics Plans
Washington, D. C. 20350

Dr. Andrew Rechnitzer
Office of the Chief of Naval Operations, OP 952F
Naval Oceanography Division
Washington, D. C. 20350

Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Department
Code N-71
Naval Training Equipment Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation Group
Orlando, FL 32813

CDR Norman E. Lane
Code N-7A
Naval Training Equipment Center
Orlando, FL 32813

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Mr. H. Talkington
Ocean Engineering Department
Naval Ocean Systems Center
San Diego, CA 92152

<u>Department of the Navy</u>

Mr. Paul Heckman
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Ross Pepper
Naval Ocean Systems Center
Hawaii Laboratory
P. O. Box 997
Kailua, HI 96734

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D. C. 20380

Dr. L. Chmura
Naval Research Laboratory
Code 7592
Computer Sciences & Systems
Washington, D. C. 20375

HQS, U. S. Marine Corps
ATTN: CCA40 (Major Pennell)
Washington, D. C. 20380

Commanding Officer
MCTSSA
Marine Corps Base
Camp Pendleton, CA 92055

Chief, C^3 Division
Development Center
MCDEC
Quantico, VA 22134

Human Factors Technology Administrator
Office of Naval Technology
Code MAT 0722
800 N. Quincy Street
Arlington, VA 22217

Commander
Naval Air Systems Command
Human Factors Programs
NAVAIR 334A
Washington, D. C. 20361

<u>Department of the Navy</u>

Commander
Naval Air Systems Command
Crew Station Design
NAVAIR 5313
Washington, D. C. 20361

Mr. Philip Andrews
Naval Sea Systems Command
NAVSEA 03416
Washington, D. C. 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 81323
Washington, D. C. 20360

Larry Olmstead
Naval Surface Weapons Center
NSWC/DL
Code N-32
Dahlgren, VA 22448

Mr. Milon Essoglou
Naval Facilities Engineering Command
R&D Plans and Programs
Code 03T
Hoffman Building II
Alexandria, VA 22332

Capt. Robert Biersner
Naval Medical R&D Command
Code 44
Naval Medical Center
Bethesda, MD 20014

Dr. Arthur Bachrach
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD 20014

Dr. George Moeller
Human Factors Engineering Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

Department of the Navy

Head
Aerospace Psychology Department
Code L5
Naval Aerospace Medical Research Lab
Pensacola, FL 32508

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Commander, Naval Air Force,
U. S. Pacific Fleet
ATTN: Dr. James McGrath
Naval Air Station, North Island
San Diego, CA 92135

Navy Personnel Research and
Development Center
Planning & Appraisal Division
San Diego, CA 92152

Dr. Robert Blanchard
Navy Personnel Research and
Development Center
Command and Support Systems
San Diego, CA 92152

CDR J. Funaro
Human Factors Engineeing Division
Naval Air Development Center
Warminster, PA 18974

Mr. Stephen Merriman
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Mr. Jeffrey Grossman
Human Factors Branch
Code 3152
Naval Weapons Center
China Lake, CA 93555

Human Factors Engineering Branch
Code 1226
Pacific Missile Test Center
Point Mugu, CA 93042

Department of the Navy

Dean of the Academic Departments
U. S. Naval Academy
Annapolis, MD 21402

Dr. S. Schiflett
Human Factors Section
Systems Engineering Test
Directorate
U. S. Naval Air Test Center
Patuxent River, MD 20670

Human Factor Engineering Branch
Naval Ship Research and Development
Center, Annapolis Division
Annapolis, MD 21402

Mr. Harry Crisp
Code N 51
Combat Systems Department
Naval Surface Weapons Center
Dahlgren, VA 22448

Mr. John Quirk
Naval Coastal Systems Laboratory
Code 712
Panama City, FL 32401

CDR C. Hutchins
Code 55
Naval Postgraduate School
Monterey, CA 93940

Office of the Chief of Naval
Operations (OP-115)
Washington, D. C. 20350

Professor Douglas E. Hunter
Defense Intelligence College
Washington, D. C. 20374

Department of the Army

Mr. J. Barber
HQS, Department of the Army
DAPE-MBR
Washington, D. C. 20310

Department of the Navy

Dr. Edgar M. Johnson
Technical Director
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Organizations and
Systems Research Laboratory
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Technical Director
U. S. Army Human Engineering Labs
Aberdeen Proving Ground, MD 21005

Department of the Air Force

U. S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D. C. 20332

AFHRL/LRS TDC
Attn: Susan Ewing
Wright-Patterson AFB, OH 45433

Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Dr. Earl Alluisi
Chief Scientist
AFHRL/CCN
Brooks Air Force Base, TX 78235

Foreign Addressees

Dr. Daniel Kahneman
University of British Columbia
Department of Psychology
Vancouver, BC V6T 1W5
Canada

Foreign Addressees

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Technology
Establishment
Teddington, Middlesex TW11 OLN
England

Director, Human Factors Wing
Defence & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
Canada

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF England

Other Government Agencies

Defense Technical Information Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 copies)

Dr. Craig Fields
Director, System Sciences Office
Defense Advanced Research Projects
Agency
1400 Wilson Blvd.
Arlington, VA 22209

Dr. M. Montemerlo
Human Factors & Simulation
Technology, RTE-6
NASA HQS
Washington, D. C. 20546

Dr. J. Miller
Florida Institute of Oceanography
University of South Florida
St. Petersburg, FL 33701

Other Organizations

Dr. Robert R. Mackie
Human Factors Research Division
Canyon Research Group
5775 Dawson Avenue
Goleta, CA 93017

Dr. Amos Tversky
Department of Psychology
Stanford University
Stanford, CA 94305

Dr. H. McI. Parsons
Human Resources Research Office
300 N. Washington Street
Alexandria, VA 22314

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311

Professor Howard Raiffa
Graduate School of Business
  Administration
Harvard University
Boston, MA 02163

Dr. T. B. Sheridan
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

Dr. Arthur I. Siegel
Applied Psychological Services, Inc.
404 East Lancaster Street
Wayne, PA 19087

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Harry Snyder
Department of Industrial Engineering
Virginia Polytechnic Institute and
  State University
Blacksburg, VA 24061

Other Organizations

Dr. Ralph Dusek
Administrative Officer
Scientific Affairs Office
American Psychological Association
1200 17th Street, N. W.
Washington, D. C. 20036

Dr. Robert T. Hennessy
NAS - National Research Council (COHF)
2101 Constitution Avenue, N. W.
Washington, D. C. 20418

Dr. Amos Freedy
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Robert C. Williges
Department of Industrial Engineering
  and OR
Virginia Polytechnic Institute and
  State University
130 Whittemore Hall
Blacksburg, VA 24061

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, N. W.
Washington, D. C. 20036

Dr. Deborah Boehm-Davis
General Electric Company
Information Systems Programs
1755 Jefferson Davis Highway
Arlington, VA 22202

Dr. Ward Edwards
Director, Social Science Research
  Institute
University of Southern California
Los Angeles, CA 90007

Dr. Robert Fox
Department of Psychology
Vanderbilt University
Nashville, TN 37240

Other Organizations

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, OK 73069

Dr. Kenneth Hammond
Institute of Behavioral Science
University of Colorado
Boulder, CO 80309

Dr. James H. Howard, Jr.
Department of Psychology
Catholic University
Washington, D. C. 20064

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Urbana, IL 61801

Mr. Edward M. Connelly
Performance Measurement
Associates, Inc.
410 Pine Street, S. E.
Suite 300
Vienna, VA 22180

Professor Michael Athans
Room 35-406
Massachusetts Institute of
Technology
Cambridge, MA 02139

Dr. Edward R. Jones
Chief, Human Factors Engineering
McDonnell-Douglas Astronautics Co.
St. Louis Division
Box 516
St. Louis, MO 63166

Other Organizations

Dr. Babur M. Pulat
Department of Industrial Engineering
North Carolina A&T State University
Greensboro, NC 27411

Dr. Lola Lopes
Information Sciences Division
Department of Psychology
University of Wisconsin
Madison, WI 53706

Dr. A. K. Bejczy
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91125

Dr. Stanley N. Roscoe
New Mexico State University
Box 5095
Las Cruces, NM 88003

Mr. Joseph G. Wohl
Alphatech, Inc.
3 New England Executive Park
Burlington, MA 01803

Dr. Marvin Cohen
Decision Science Consortium
Suite 721
7700 Leesburg Pike
Falls Church, VA 22043

Dr. Wayne Zachary
Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090

Dr. William R. Uttal
Institute for Social Research
University of Michigan
Ann Arbor, MI 48109

Dr. William B. Rouse
School of Industrial and Systems
Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Other Organizations

Dr. Richard Pew
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02238

Dr. Hillel Einhorn
Graduate School of Business
University of Chicago
1101 E. 58th Street
Chicago, IL 60637

Dr. Douglas Towne
University of Southern California
Behavioral Technology Laboratory
3716 S. Hope Street
Los Angeles, CA 90007

Dr. David J. Getty
Bolt Beranek & Newman, Inc.
50 Moulton street
Cambridge, MA 02238

Dr. John Payne
Graduate School of Business
Administration
Duke University
Durham, NC 27706

Dr. Baruch Fischhoff
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Andrew P. Sage
School of Engineering and
Applied Science
University of Virginia
Charlottesville, VA 22901

Denise Benel
Essex Corporation
333 N. Fairfax Street
Alexandria, VA 22314