MARTIN MARIETTA ENERGY SYSTEMS LIBRARIES

3 4456 0282181 9

ORNL/TM-10898

ornl

OAK RIDGE
NATIONAL
LABORATORY

MARTIN MARIETTA

People Versus Probability

Bruce Tonn
Richard Goeltz

OAK RIDGE NATIONAL LABORATORY
CENTRAL RESEARCH LIBRARY
CIRCULATION SECTION
4500N ROOM 175
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON
If you wish someone else to see this report, send in name with report and the library will arrange a loan.

OPERATED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161
NTIS price codes—Printed Copy: A03; Microfiche: A01

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/TM-10898

PEOPLE VERSUS PROBABILITY

BRUCE TONN

RICHARD GOELTZ

Energy Division

August 1988

Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
operated by
Martin Marietta Energy Systems, Inc.
under contract No. DE-AC05-840R21400
for the
U.S. Department of Energy

MARTIN MARIETTA ENERGY SYSTEMS LIBRARIES

3 4456 0282181 9

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
ABSTRACT	vii
PEOPLE VERSUS PROBABILITY	1
EXPERIMENT 1: Probabilistic Context	5
SUBJECTS	5
METHOD	5
RESULTS	8
EXPERIMENT 2: Representative Context	15
SUBJECTS	15
METHOD	15
RESULTS	17
DISCUSSION	23
REFERENCES	27

LIST OF TABLES

		Page
Table 1.	Means of the likelihoods of events happening on any given day (N=21)	6
Table 2.	Definitions of six frequentistic rules	9
Table 3.	Means of observed Z values and predicted Z values for six rules over 10 cases for the frequentistic data	11
Table 4.	Cluster analysis of frequentistic data	13
Table 5.	Means for the likelihoods of personality characteristics relating to a good personality (N=21)	16
Table 6.	Definitions of six evidential rules	18
Table 7.	Relationships between observed Z values and predicted Z values for six rules over four cases for the evidential data	20
Table 8.	Cluster analysis of evidential data	22

ABSTRACT

This research explores how subjects combine estimates of uncertainty in probabilistic and evidential contexts. Two major computer programs written in Common Lisp asked subjects questions about the likelihoods and conjunctions of independent events. The results suggest that in the probabilistic context the best model to describe individual decision making is not the product rule but a minimum rule, and models varied as the magnitude of the likelihoods varied. In the evidential context, subjects appeared to use a maximum rule, although some evidence supports the use of the certainty factor rule. Subjects had difficulty in combining contradictory evidence.

PEOPLE VERSUS PROBABILITY

Numerous psychological studies have shown that humans suffer from systematic biases in decision making. One of the more interesting problems is known as the "conjunction effect." Subjects are typically asked to rank order the likelihood of two single events, such as A and B, and their conjunction, A & B. Probability theory maintains that $P(A) \leq P(A\&B)$ and $P(B) \leq P(A\&B)$. However, Tversky and Kahneman (1983) found that under certain circumstances, individuals will assert that $P(A) > P(A\&B) > P(B)$. This finding held in numerous instances, from personality perception to medical judgment, from sporting to forecasting.

Several attempts have been made to explain conjunction effects. Lack of statistical training is only a partial explanation because numerous subjects with training defended their answers (Tversky and Kahneman 1983). Another explanation is that subjects intentionally apply the representativeness heuristic, which yields "an assessment of the degree of correspondence between a sample and a population" (Tversky and Kahneman 1983). This explanation intuitively appears related to subjects ranking as more likely a person, who was described as having the characteristics of a feminist, as being a bank teller and a feminist rather than just a bank teller (Tversky and Kahneman 1982).

This line of research has, in addition to identifying conjunctive effects, analyzed rules or heuristics people may apply to conjunctive problems. For example, Wyer (1976) found that multiplying and averaging models, especially when a high and low probability of events are being conjoined, fit the data well. Yates and Carlson (1986) found similar results and also hypothesized that subjects may have applied a "signed sum" model. In this model, an event's likelihood is represented by a

positive number, a negative number, or zero, depending on whether the event is considered likely, unlikely, or neither, respectively. The likelihood value for a conjunction is the sum of those for the constituent marginal events" (Yates and Carlson 1986, P230).

This article contributes to the literature on conjunctive effects in three ways. Most importantly, the two experiments reported herein are designed to highlight how a decision context frames conjunctive effects. Two decision contexts are explored. The first, termed probabilistic, refers to conjunctive problems that individuals may solve by using data on the frequency of events. The second, termed evidential, refers to problems solvable by making similarity comparisons between evidence and a possible diagnosis/decision. Physicians, it can be argued, may use evidential reasoning to compare a patient's symptoms to a possible disease. Lawyers may also utilize evidential reasoning when comparing a current case to those providing precedent. Gamblers, actuaries, and schedulers, for example, should use probabilistic reasoning.

One hypothesis being investigated is that heuristics used by subjects in probabilistic and evidential reasoning are different from each other. Another hypothesis is that subjects do not apply normatively correct rules in either context. Instead, it is argued that individuals use cognitively simpler rules.

This research delves deeper into the nature of decision making than past work on conjunctive effects because the two contexts represent fundamentally distinct approaches to reasoning, not just a compilation of interesting problem topics. The results indicate that subjects indeed solve conjunctive problems differently in each context. Thus, descriptively speaking, it has been established that subjects can be induced to

reason differently through deliberate changes in problem contexts. Also models that best describe the data do not correspond to models that have been argued to be normatively correct for probabilistic and evidential reasoning. While it is the intention of this paper to argue that it is important to distinguish between probabilistic and evidential reasoning and assume that normative rules appropriate for one context are not necessarily appropriate for the other, this paper does not attempt to advocate any particular normative approach for reasoning in either context.

A second contribution relates to the novelty of the experiments. First, subjects were allowed some flexibility in representing probability estimates. Second, in both experiments, a computer survey tool was able to individualize questions, allow answer editing, and dynamically query subjects about possible inconsistencies in their answers. Thus, the database, in all likelihood, has relatively high reliability. More sophisticated software, possibly involving machine learning, could expand on these themes to yield more interesting and complex data sets.

Lastly, in the discussion section, a link between research on conjunctive effects and expert systems is established. This could be an important relationship because builders of expert systems have not been cognizant of the psychology of decision making under uncertainty and psychologists may find many interesting problems to explore related to expert systems and human expert reasoning.

EXPERIMENT 1: Probabilistic Context

SUBJECTS

Twenty-one subjects completed two 30-minute sessions on the computer. Nineteen subjects were members of the technical staff of Oak Ridge National Laboratory; the other two also had technical backgrounds. The subjects were self-selected based on their desire to contribute to this research and on their time constraints.

METHOD

The survey instrument consisted of two main computer programs written in Common Lisp on a VAX computer. One program controlled the first session; the other program controlled the second session. An introductory message that explained the goals of the research preceded the first session. The experimenter provided no formal introduction or instructions on how to answer first session questions. In the first session, each subject was asked 31 questions pertaining to "the likelihood of events happening on an average day in their lives." The set of events was designed to contain high, medium, and low probability events, such as "What is the likelihood of you taking a shower or bath," "going to the grocery store," and "being hit by an object falling from the sky," respectively (Table 1). At the end of session 1, the subjects were allowed to review and change their answers. Each likelihood could be represented as a probability (e.g., PR 0.8) or a chance (e.g., CH 1 in 10). A help facility which provided information on representing likelihoods was available.

The second session was administered at least one day after the first. Subjects were presented with questions that inquired about the

Table 1. Means of the likelihoods of events happening on any given day (N=21)

No.	Question Description	Mean
1	That it will rain or snow	.33
2	That you will be in an automobile accident	.016
3	That you will receive in the mail an incorrect bank statement	.011
4	That your television will break	.046
5	That you will go to a grocery store	.31
6	That you will have a good day at school or work	.67
7	That you will receive a phone call from friends or relatives	.46
8	That you will be awakened at night by outside noises	.14
9	That you will have a sandwich for lunch	.31
10	That you will listen to the radio	.84
11	That you will throw your garbage out	.45
12	That you will eat dinner out	.17
13	That you will get some exercise	.58
14	That you will read a daily newspaper	.63
15	That you will write a letter to your parents	.022
16	That you will receive a phone call asking for money	.065
17	That you will take a shower or bath	.96
18	That you will be hit by an object falling from the sky	.00065
19	That you will suffer from a headache	.078
20	That you or someone you live with will wash some clothes	.44
21	That you will watch a sports show on television	.19
22	That you will floss your teeth	.56
23	That you or someone you live with will use the oven	.57
24	That you will watch a news program on TV	.55
25	That you will drink at least one glass of milk	.52
26	That you or someone you live with will vacuum your home	.28
27	That you will go to a movie theater	.094
28	That you will write a check	.39
29	That you will drink at least one beer	.16
30	That you will replace a light bulb in your home	.10
31	That you will take a nap	.12

likelihood of compound events (e.g., "What is the likelihood of your being awakened at night by outside noises and having a sandwich for lunch."). The experimenter instructed the subjects to consider the events as unrelated, for the purposes of the study. Previous answers

were not supplied and answers were not allowed to be reviewed. However, the program did perform analysis on the data to determine whether answers were inconsistent.

To understand how this was done, we must describe how the second session program worked. Specifically, the program had the goal of filling a 6x6 array, where the X and Y coordinates represented points on the likelihood continuum from 0.0 to 1.0. The purpose was to find single likelihood answers that fit into each of the six "compartments" and then put them together into combination questions in order to determine whether or not conjunctive effects were dependent on the magnitudes or order of the likelihoods being combined. Inconsistencies arise when a subject's array does not exhibit a monotonically increasing pattern from the 0.0,0.0 point to the 1.0,1.0 point. For example, let X and Y represent subject estimated likelihoods for events A and B. Let Z represent the subject-estimated likelihood for the event A&B. The Zs should monotonically increase as the Xs and Ys increase; otherwise there are logical inconsistencies in the data.

If an inconsistency was found, then the program would ask the subject if he or she agreed with the answer to the combination question. If yes, then the program would search for additional inconsistencies. If no, then the program would ask for agreement on the two single likelihood estimates that comprised the question. If either of these answers was no, then the program would ask the single likelihood question(s) again. All cells in the array that were dependent on the now-changed single likelihood estimates were nulled out and the program would go about asking combination questions to refill the cells with valid data.

RESULTS

It was hypothesized, a priori, that subjects would not exhibit behavior consistent with normative probabilistic reasoning. Further, it was hypothesized that deviations from the probability model could be explained by subjects applying heuristics that economize on mental activity. By mental economy, we mean that subjects would minimize the number of mental operations required to produce answers. The easiest mental operation would be direct access to memory, as opposed to active problem solving (Kahneman and Tversky 198[4]). Within the active problem solving context, multiplying two single digit numbers is easier then multiplying two double digit numbers. Geary et al. (1986) estimate that multiplication difficulty increases as the product of the number digits being multiplied.

Only the easiest conjunctive problems were hypothesized to be solved using methods resembling the product rule. Specifically, the easiest conjunctive problems involve one digit probabilities with straightforward relationships to memories of multiplication tables. This qualification has a tendency to limit product rule applicability to conjunctive problems involving high and medium probabilities represented in single digit form.

All conjunctive problems may be simplified by using averaging, anchor-adjustment, and min/max types of rules. In the probabilistic context, an anchor-adjustment heuristic entails choosing one of the two likelihoods (e.g., the lowest) being conjoined as an anchor and then adjusting it downward depending on the magnitude of the other likelihood. Applying a min rule may result in a more biased conjunction but is simpler cognitively. An average rule is least appropriate normatively. A priori, we

hypothesized that people may use these kinds of heuristics when faced with combining low probabilities.

Table 1 presents the means associated with likelihoods of the 31 events. The range of likelihoods, in the aggregate, is broad, as was desired. The ranges were broad enough for each subject to allow the second session computer program to ask the full range of conjunction questions.

To assist in understanding the types of processes that subjects may have applied in solving the conjunctive problems, observations were compared to six specific rules (Table 2). X and Y are the likelihoods to be conjoined and Z is the conjunction. The six rules are arranged in ascending order with respect to outputting Z values. Rule 2 is the product rule. Rule 1 is the product rule squared and was included to capture the possibility that subjects could apply heuristics that greatly overemphasize the uncertainty associated with the happening of two independent events.

Table 2. Definitions of six frequentistic rules

Rule	Label	Formula
1	Product rule squared	$(X * Y)^2 = Z$
2	Product rule	$X * Y = Z$
3	Anchor adjustment	$[X*Y + MIN(X,Y)]/2 = Z$
4	Minimum	$MIN(X,Y) = Z$
5	Average	$(X + Y)/2 = Z$
6	Certainty factor	$X + (1 - X) * Y = Z$

Rule 3 is labeled the anchor-adjustment rule. As defined, it is the unweighted average of the result of the product rule and the minimum rule (#4). It represents an attempt, however crude, to capture instances where subjects may estimate a conjunction by adjusting the minimum probability by an amount commensurate with the product of X and Y. Rule 5 represents an unweighted average of X and Y. Rule 6, the certainty factor rule, was proposed by Shortliffe (1976) as a way to combine uncertainties in expert systems and a priori, was deemed more applicable for evidential reasoning. It was included in this experiment to demonstrate its apparent unappropriateness for probabilistic reasoning.

Table 3 presents means of observed Z values and Z values calculated using the six rules and subject-supplied Xs and Ys. Column 1 contains the means for all conjunction questions put to subjects (21 x 36 data points). The other columns contain means subset on the magnitudes of the Xs and Ys. HIGH is defined as likelihood greater than 0.45, MED as between 0.05 and 0.45, and LOW as below or equal to 0.05. The ranges were set to split the range of observed X and Y values into thirds.

The results in Table 3 support the general hypothesis of mental economy. In the aggregate (Col. 1), the observed mean falls between Rules 3 and 4, the anchoring and the minimum rules. Only when subjects combined high likelihoods (Col. 2) did the product rule appear to be used. In cases where either X or Y is a low likelihood (Cols. 3, 5, 7 and 9) the rule with the closest prediction is the minimum rule. This is a fairly robust observation because the observed means are essentially equal in these cases. A minimum type of rule appears to have also been applied to the low/low cases. Thus in most cases, the data suggest that subjects used a cognitively simple rule to conjoin likelihoods.

Table 3. Means of observed Z values and predicted Z values for six rules over 10 cases for the frequentistic data.*

Rule	1 All	2 High/High	3 High/Low	4 High/Med	5 Low/High	6 Low/Low	7 Low/Med	8 Med/High	9 Med/Low	10 Med/Med
Z Observed	.134	.50	.025	.229	.023	.0044	.018	.254	.023	.126
1. Product rule squared	.050 (.64)	.374 (.48)	.00016 (.16)	.041 (.33)	.00019 (.47)	$1.39*10^{-7}$ (.34)	$2.3*10^{-5}$ (.17)	.040 (.53)	$1.5*10^{-5}$ (.14)	.0041 (.40)
2. Product rule	.110 (.75)	.572 (.46)	.0062 (.22)	.175 (.38)	.0064 (.53)	$9.7*10^{-5}$ (.35)	.0022 (.32)	.174 (.54)	.0019 (.22)	.054 (.45)
3. Anchor adjustment	.127 (.76)	.606 (.46)	.0073 (.22)	.203 (.36)	.0071 (.55)	.0015 (.38)	.0056 (.36)	.202 (.52)	.0057 (.24)	.111 (.39)
4. Minimum	.145 (.76)	.640 (.44)	.0084 (.23)	.232 (.32)	.0078 (.56)	.0029 (.38)	.0089 (.37)	.231 (.47)	.0095 (.23)	.168 (.36)
5. Average	.334 (.66)	.756 (.42)	.375 (.09)	.491 (.37)	.380 (.15)	.0083 (.30)	.118 (.11)	.490 (.45)	.117 (.06)	.236 (.45)
6. Certainty factor	.557 (.49)	.971 (.23)	.744 (.08)	.806 (.31)	.753 (.12)	.017 (.30)	.235 (.10)	.806 (.32)	.233 (.05)	.419 (.44)

*Numbers in parentheses are correlation coefficients between the observed Z values and the particular predicted Z values.

One of the strengths of using a computer to gather data is that it can gather and organize large amounts of data from each individual. Thus, it facilitates data analysis aimed at determining what any one individual may be doing with respect to combining uncertainty heuristics. Given that the second session software could not automatically learn individual's combining heuristics, the next best approach is to cluster individuals by the kinds of rules they may use. To facilitate cluster analysis, conjunctions were aggregated into just three categories, a high/high category, one for all the combinations with a medium likelihood but no low likelihood, and one for all combinations with a low likelihood.

Next, for each observation (Z value) for each subject within each category, a scale was developed that places the observation in a space defined by the Z value outputs for the six rules. Thus, an observed z that is exactly equal to the prediction of Rule 2 would be valued at 2.0. If the observed z falls close to Rule 2 but between Rules 2 and 3, then the observation may be valued at 2.2. In this way, the kinds of rules that the subjects may be using to answer any question in the three categories of combinations may be characterized.

Next, a cluster analysis, using Ward's method, was performed over the means for the characterizations of the observations using variables defined over the three categories (Table 4). The results indicate that there is wide variation across individuals in heuristics used to combine uncertainty estimates in the probabilistic context. Over four clusters, subjects grouped into Cluster 2 (N=5) came closest to making decisions similar to using the probability rule, where no characteristic mean exceeds 2.5. Interestingly, the high/high category mean is very low, almost to Rule 1.

Table 4. Cluster Analysis of Frequentistic Data*

Cases of Likelihood Combinations	Means Cluster			
	1 (N=9)	2 (N=5)	3 (N=3)	4 (N=3)
1. Low-high, low-low, low-med high-low, med-low	3.54	2.50	2.65	4.13
2. Med-med, high-med, med-high	3.21	2.38	2.61	4.00
3. High-high	3.83	1.1	2.69	1.22

*Ward's method was used. R^2 for four clusters = .78. The values in the table refer to the mean of the observed Z values for each cluster over each case when placed on a six point scale defined over the rules listed in Table 2.

In contrast, cluster 4 subjects (N=3) exhibit Rule 2 behavior in the high/high category but exhibit Rule 4 behavior in the other two categories. These subjects appear to be using very different heuristics conditional on the magnitudes of the likelihoods being combined. Subjects in clusters 1 (N=9) and 3 (N=3) exhibit consistent behavior over the three categories; cluster 1 subjects lean toward rules between Rules 3 and 4, whereas cluster 3 subjects lean toward rules between Rules 2 and 3.

In summary of Experiment 1, the data suggest that individuals do not use the product rule to combine estimates of uncertainty in the frequentistic context. Instead, cognitively simpler processes appear to have been employed. The data also suggest that the heuristics used are extremely idiosyncratic across individuals and dependent on the magnitudes of the estimates being combined.

EXPERIMENT 2: Representative Context

SUBJECTS

Twenty-one subjects completed two 30-minute sessions on the computer. Nineteen had already completed Experiment 1. The two new subjects possessed technical backgrounds.

METHOD

This experiment was structured exactly like the first. In the first session, each subject was asked to estimate the likelihood of a person having a GOOD personality that possessed a certain personality characteristic. Seventy-five characteristics were given (Table 5), approximately equally divided between positive, neutral, and negative characteristics. In the second session, the subjects were presented with conjunctive questions (e.g., what is the likelihood of a person having a GOOD personality that is "loyal" and "open-minded"). In this experiment, the experimenter did not instruct subjects to treat the personality characteristics as independent. The software endeavored, as in the first experiment, to present conjunctive questions that paired a diversity of characteristics.

Answers were restricted to only one representation, that being a number in the range -1.0 to 1.0. For positive characteristics, the answers should have been in the >0.0 to 1.0 range, for negative characteristics in the <0.0 to -1.0 range, and for completely neutral characteristics equal to 0.0.

**Table 5. Means for the likelihoods of personality characteristics relating to a good personality*
(N=21)**

No.	Personality Characteristics	Mean	No.	Personality Characteristics	Mean
1	Open minded	.75	39	Time waster	-.24
2	Curious	.51	40	Impulsive	0
3	Intelligent	.34	41	Unloyal	-.71
4	Has sense of humor	.75	42	Unprincipled	-.70
5	Generous	.66	43	Passive	-.15
6	Youthful	.31	44	Impatient	-.38
7	Caring of others	.80	45	Nervous	-.30
8	Modest	.33	46	Unconfident	-.35
9	Trusting	.52	47	Avoids challenges	-.40
10	Fun loving	.75	48	Stubborn	-.37
11	Idealistic	.34	49	Opportunistic	-.12
12	Polite	.62	50	Does not look to future	-.10
13	Brave	.34	51	Unconventional	.16
14	Efficient	.22	52	Conventional	.06
15	Cautious	.06	53	Average intelligence	.16
16	Loyal	.60	54	Clean	.39
17	Honest	.71	55	Enjoys sex	.41
18	Active	.46	56	Needs love	.07
19	Patient	.62	57	Tall	.03
20	Calm	.38	58	Short	-.02
21	Self-confident	.50	59	Blond	-.01
22	Enjoys challenges	.43	60	Wears glasses	.02
23	Steadfast	.47	61	Follows the crowd	-.12
24	Purposeful	.34	62	Athletic	.20
25	Forward looking	.42	63	Heavy	0
26	Opinionated	-.18	64	Light	.10
27	Uninterested in things	-.64	65	Brown-eyed	0
28	Underachiever in school	-.20	66	Messy	-.11
29	Boring	-.72	67	Smokes	-.13
30	Stingy	-.56	68	Workaholic	-.33
31	Childish	-.47	69	Likes to read	.34
32	Egocentric	-.55	70	Energetic	.64
33	Conceited	-.68	71	Healthy	.22
34	Gullible	-.10	72	Sickly	-.20
35	Shy	-.11	73	Talkative	.26
36	Pessimistic	-.63	74	Responsible	.69
37	Rude	-.84	75	Handsome	.06
38	Insensitive	-.70			

*Characteristics 1-25 were designed to be primarily positive, 26-50 negative, and 51-75 neutral.

RESULTS

Experiment 2 had two purposes, to explore further the mental economy hypothesis and to discover whether the evidential context has a different influence upon conjunctive effects than the probabilistic context. The context of this experiment was deliberately chosen to try to elicit a representativeness type of response from the subjects. That is, given one or two personality characteristics, subjects are hypothesized make a similarity comparison rather than a probabilistic analysis on the likelihood of having a GOOD personality.

Conjunctive questions concerning similarity comparisons are hypothesized to be answered using heuristics different from those utilized to answer probabilistic conjunctions. Also, similarity comparisons necessarily assume dependence among characteristics, symptoms, etc., so that the independence assumption, which was important for Experiment 1, is not needed here. Unfortunately, there is no accepted norm concerning similarity conjunctions. However, the literature does reveal possibilities.

The most interesting possible norm is known as certainty factor theory (Shortliffe 1976). Certainty factors were developed to represent uncertainty in expert systems. Several constraints in probability theory lead to their development. For this research, the most important constraint is that experts did not want uncommitted probability mass (e.g., mass left over from an estimate that symptoms X, Y and Z relate to disease S with a likelihood of 0.80) to be distributed to support the likelihood of other possible disease. In essence, the experts wanted a measure for similarity comparisons, not one to establish probability distributions.

Rule 3 in Table 6 is the certainty factor rule. It is very different from the product rule; given an X and Y, the certainty factor rule will output a conjunction (Z-value) that is greater than or equal to the highest term, whereas the product rule will output a value less than or

Table 6. Definitions of six evidential rules

Rule No.	Label	
1. Relative Minimum	IF X and Y are both positive,	Z = MIN(X,Y)
	IF X and Y are both negative,	Z = MAX(X,Y)
	Otherwise	Z = Minimum of the absolute values of X and Y
2. Relative Maximum	IF X and Y are both positive,	Z = MAX(X,Y)
	IF X and Y are both negative,	Z = MIN(X,Y)
	Otherwise	Z = Minimum of the absolute values of X and Y
3. Certainty Factor	IF X and Y are both positive,	Z = X + (1-X) * Y
	IF X and Y are both negative,	Z = -[-X+(1+X) * -Y]
	Otherwise	Z = X + Y
4. Ultra-Certainty Factor	IF X and Y are both positive,	Z = 1-[0.5 *(1-X) *(1-Y)]
	If X and Y are both negative,	Z = -1 + [0.5 *(1+X) *(1+Y)]
	Otherwise, if ABS(X) is greater than ABS(Y)	Z = X*(1±Y)/(1+X*Y) ± 0.5 * (1±X)(1±Y)/[1+(X*Y)]
	Otherwise	Z =Y*(1±X)/(1+X*Y)± 0.5 * (1±X)(1±Y)/[1+(X+Y)]
5. Product Rule	Usually	Z = X * Y
	BUT if X and Y are both negative	Z = -(X * Y)
6. Additive	Usually,	Z = X + Y
	BUT, if Z is greater than 1	Z = 1
	BUT, if Z is less than -1	Z = -1

equal to the lowest term. In effect, the certainty factor rule accumulates likelihood mass and will approach 1.0 as the evidence gets stronger. Although other alternatives to probability theory have gained some prominence in the expert systems area (e.g., fuzzy sets, belief functions), certainty factors are by far the most widely used.

With respect to mental economy, the certainty factor rule is relatively taxing because it requires both multiplicative and additive operations. Simpler processes that might be observed that are consistent with certainty factor/similarity comparison type answers would be maximum rule or additive rule (Rules 2 and 6, Table 6, respectively). We would not expect the data to support processes resembling the product rule or minimum rule (Rules 1 and 5, Table 6, respectively) because these rules do not act to accumulate evidence.

Table 5 contains the means of the answers collected in session 1 from the 21 subjects. The range of answers was sufficient for the software to construct for all subjects conjunctive questions containing diverse personality characteristics. However, two subjects did not give answers with enough diversity for the software to construct a full 36 questions for the second session.

Table 7 summarizes the results in aggregate form. Analysis was subset into four cases, where the conjunctive characteristics are <u>strongly supportive</u> of each other, <u>weakly supportive</u>, <u>strongly contradictory</u>, and <u>weakly contradictory</u>. Figure 1 depicts how the assignments were made. We hypothesized that subjects would provide more consistent answers when characteristics are supportive of either a positive or negative answer than when characteristics are in some sense opposite of each other.

Table 7. Relationships between observed Z values and predicted Z values for six rules over four cases for the evidential data* (% of responses)

Case		Rules 1 Relative minimum	2 Relative maximum	3 Certainty factor	4 Ultra-certainty factor	5 Product rule	6 Additive rule
Strongly supportive (N=192)	Over-predictive	9	50	60	70	4	69
	Hit	14	29	24	20	9	21
	Under-predictive	77	21	16	10	87	10
Weakly supportive (N=132)	Over-predictive	6	22	45	84	0	48
	Hit	21	33	20	2	15	20
	Under-predictive	73	45	35	14	85	32
Strongly Contra-dictory (N=148)	Over-predictive	20	57	24	38	12	24
	Hit	6	7	12	18	9	12
	Under-predictive	74	36	64	44	79	64
Weakly contra-dictory (N=165)	Over-predictive	4	51	42	38	0	43
	Hit	10	8	13	8	10	13
	Under-predictive	85	41	44	54	90	44

*A hit occurs when the predicted value is ± 5% of the observed value. A rule is overpredictive if the predicted value is greater than 5% of the observed value. A rule is underpredictive if the opposite is true.

A rule hits an observation when its prediction falls within + or -5% of the observed value. A rule is overpredictive when its prediction is greater than 5% of the observation and underpredictive when its prediction is less than 5% of the observation. A good fit with a rule exists when there are a large number of hits and an even distribution

between the over- and underpredictive cases. For the strongly supportive cases, the best fit appears to be Rule 2, the relative maximum rule, followed by Rule 3, the certainty factor rule. The minimum rule (1) and the product rule (5) significantly underpredict the observations, as hypothesized, whereas the addition rule (6) and the ultra Certainty Factor (CF) rule (4) overpredict. Even the two best fit rules, 3 and 4, tend to be too overpredictive.

The results are similar in the weakly supporting case. Rules 1 and 5 underpredict and Rule 4 overpredicts. Rule 2 fits the data much better than in the strongly supportive case and Rules 3 and 6 are better balanced. The data suggest that, in the aggregate, the relative maximum rule best models the subjects' heuristics in solving conjunctions involving similarity comparisons.

The results change a little in the contradictory cases. In the strongly contradictory case, none of the rules perform as well, from observing the percentage of hits. Rule 4 has the highest number of hits and good balance. Except for the certainty factor rule (3), the others have both a low number of hits and unbalanced ranges. Rules 3 and 6 tie for best in the weakly contradictory case. These results indicate much diversity in answers given contradictory personality characteristics, with a possible strong bias toward the positive or negative in the strongly contradictory case. In general, it appears that the certainty factor rule is as good as any in describing decision making using contradictory evidence.

Table 8 presents results of a cluster analysis. The clusters were developed in a manner different from Experiment 1. In this case, the subjects were clustered over four variables, defined as the mean difference

Table 8. Cluster analysis of evidential data*

Case	Clusters 1 (N=13)	2 (N=5)	3 (N=2)	4 (N=1)
Strongly supportive (Anchor Rule 2)	0.05	-0.01	-0.02	-0.01
Weakly supportive (Anchor Rule 2)	-0.03	-0.12	0.03	-0.40
Strongly contradictory (Anchor Rule 4)	-0.06	-0.11	-0.28	-0.40
Weakly contradictory (Anchor Rule 3)	0.00	-0.18	-0.11	-0.37

*Used Ward's method. R^2 for four clusters = 0.77. The anchor rule numbers refer to the rules in Table 7. The values in the table refer to the mean difference between the observed Z value and the anchor rule's predicted Z value for all subjects in a cluster over all questions in a case.

between the predicted z of best fit rule from table 7 and the observed z, for each of the four cases. Thus, Rule 2 was used as the best fit rule for the strongly and weakly supportive cases and Rules 4 and 3 were used in the strongly and weakly contradictory cases, respectively.

The results indicate that the majority of the subjects (cluster 1, N=13) followed the pattern of using Rules 2, 4, and 3 as noted above. Five out of twenty-one (cluster 2) combined their estimates of uncertainty more strongly than the rules predicted in all cases except the strongly supportive case. Clusters 3 and 4 contain subjects who combined their estimates much more strongly in the contradictory cases. Thus, as seen in Experiment 1, there is diversity among subjects in how they solved combination problems.

DISCUSSION

Research reported herein demonstrated four major points. One, the subjects used different heuristics to combine estimates of uncertainty between the probabilistic and representativeness contexts. Broadly, the best models used in the former context resemble a minimum rule and some version of the product rule whereas the best models in the latter context resemble a maximum rule and the certainty factor rule.

Two, the results suggest that in the probabilistic context the subjects did not widely use the expected heuristic, the product rule. Instead, the data indicate that, in the aggregate, the subjects used cognitively simpler heuristics, such as a minimum rule. Answers to conjunctive questions were very dependent upon the magnitudes of the likelihoods being combined. An extreme version of the product rule describes heuristics used when both probability estimates are relatively high. A minimum rule describes heuristics used when at least one of the likelihoods is low. On a disaggregated level, the data indicate that individuals apply heuristics very idiosyncratically. A cluster analysis identified groups of individuals who appear to use something akin to the product rule, another which uses the minimum rule, and yet another which indicates that heuristics vary widely given different likelihood magnitudes.

Three, the data cast doubt on the wide use of the certainty factor rule for combining estimates in the evidential context. The cognitively simpler maximum rule best describes the heuristics used when the two estimates support each other. This finding is particularly strong across individuals faced with combining two strongly supportive pieces of

evidence. On the other hand, the subjects had difficulty with contradictory pieces of evidence. Subjects basically differed only on how they combined contradictory evidence.

Four, this research demonstrated the utility of using the computer to collect data associated with decision making under uncertainty. The programs allowed the subjects to review answers and also probed the consistency of the subject's answers. Thus, in some sense the data may be of higher quality than obtainable from pencil and paper studies. Also, the programs tailor-made the second sessions for each subject, conditional on the answers in the first session. This flexibility allowed the collection of data not possible with paper and pencil tests.

Future research should focus on three areas: improving the software, exploring more complex decision problems, and linking research in conjunctive effects to expert systems research. More work can be done to improve the computer programs. Most important would be the development of a true machine learning algorithm that could model the heuristics used by each subject. The algorithm would have to create and test hypotheses, using its ability to quickly manipulate data.

More complex problems would explicitly incorporate time periods and explore what rules subjects may apply to the probabilistic context when dependencies exist between events. Problems should also begin to incorporate gains and losses so that data can be collected to support or reject the various models of economic decision making under uncertainty that have been proposed in the literature. Better understanding about how individuals combine uncertainties will allow researchers to separate uncertainty biases from value biases. Results of such research should help in training people to make better decisions under uncertainty.

Lastly, this line of research has potentially important implications for the artificial intelligence subfield of expert systems. Expert systems are computer programs which are designed to make decision like human experts. Most work has been done in the medical arena, where expert systems yield medical diagnoses based on information provided by system users. MYCIN, which diagnoses infectious diseases, is probably the most widely known expert system (Buchanan and Shortliffe 1984). Other expert systems analyze mass spectrometer data (Buchanan and Feigenbaum 1979), search for valuable mineral deposits (Duda et al. 1976), and configure VAX computer systems (McDermott 1982). Waterman (1985) and Buchanan (1986) provide comprehensive reviews of expert system research and development.

Expert systems are typically composed of a knowledge base and an inference engine. The former usually consists of a set of if-then rules which carry an estimate of uncertainty that relates how confident the expert is that the "if" part actually relates to the "then" part. Using the medical example, the estimate of uncertainty might relate symptoms to a diagnosis. The inference engine processes the rules and the accompanying estimates of uncertainty. Thus, the inference engine needs explicit rules about how to combine estimates of uncertainty.

The AI literature contains numerous proposals concerning alternative combination methods. Many researchers still defend the probability/Bayesian approach (e.g., Cheeseman 1985). Others have found probability wanting because it is: (1) impractical because data do not exist or experts do not have the time to estimate all combinations of possible conditional probabilities (Quinlan 1986); and (2) theoretically deficient

because it does not handle, in a logical way, uncommitted probability mass (Shafer 1976).

Thus, AI researchers have been working on finding a replacement for probability theory and have developed methodologies that have names such as certainty factors (Shortliffe 1976), endorsements (Cohen 1985), Dempter-Shafer theory (Shafer 1976; Shortliffe and Buchanan 1984), and fuzzy set theory (Negotia 1985, and Zadeh 1985). Each has strengths and weaknesses, methodologically speaking, but all suffer from a total lack of confirmation via empirical psychological research.

Research reported herein has demonstrated that the application of heuristics is context dependent, a notion not yet entertained in the expert systems field. The data also suggest that two paradigms pushed in the expert systems field, probability and certainty factors, may be unvalidated, psychologically speaking. The data also indicate differences across subjects; thus expert systems may need their combining rules to be tailored to each expert.

Psychological research can contribute to expert systems research by better identifying combining heuristics and inferential heuristics (i.e., if A causes B and B causes C, what is the relationship between A and C?) Also, expert system developers need psychologically validated methods of eliciting from experts their uncertainty heuristics.

REFERENCES

Buchanan, B., 1986. "Expert Systems: Working Systems and the Research Literature," Expert Systems, 3(1), 32-54.

Buchanan, B. and Feigenbaum, E., 1979. "DENDRAL and META-DENDRAL: Their Applications Dimension," Artificial Intelligence, 11, 5-24.

Buchanan, B. and Shortliffe, E., 1984. Rule-Based Expert Systems, Addison-Wesley, Reading, Mass.

Cheeseman, P., 1985. "In Defense of Probability," Proceedings of the Ninth International Joint Conference on Artificial Intelligence, 1002-1009.

Cohen, P., 1985. Heuristic Reasoning About Uncertainty: An Artificial Intelligence Approach, Pitman, Boston, Mass.

Duda, R., Gasching, J. and Hart, P., 1976. "Model Design in the PROSPECTOR Consultant System for Mineral Exploration," in Expert Systems in the Microelectronic Age, D. Michie (ed.), Edinburgh University Press, Edinburgh, 153-167.

Geary, D., Widman, K., and Little, T., 1986. "Cognitive Addition and Multiplication: Evidence for a Single Memory Network" Memory and Cognition 14, 478-487.

Kahneman, D. and Tversky, A., 1984. "Choices, Values, and Frames," American Psychologist, 39, 341-250.

McDermott, J., 1982. "R1, A Rule-Based Configurer of Computer Systems," Artificial Intelligence, 19, 38-88.

Negotia, C., 1985. Expert Systems and Fuzzy Systems, Benjamin/Cummings, Menlo Park, Calif.

Quinlan, R., 1986. "INFERNO: A Cautious Approach to Uncertain Inference," in Expert Systems: Techniques, Tools, and Applications, Klahr, P. and Waterman, D. (eds), Addison-Wesley, Reading, Mass.

Shafer, G., 1976. A Mathematical Theory of Evidence, Princeton University Press, Princeton.

Shortliffe, E., 1976. Computer-Based Medical Consultations: MYCIN, Elsevier, New York.

Shortliffe, E. and Buchanan, B., 1984. "A Model of Inexact Reasoning in Medicine," in Rule-Based Expert Systems, Buchanan, B. and Shortliffe, E. (eds), Addison-Wesley.

Tversky, A. and Kahneman, D., 1983. "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," Psychological Review, 90 (4), 293-315.

Tversky, A. and Kahneman, D., 1982. "Judgements of and by Representativeness," pp. 84-100 in Judgement Under Uncertainty: Heuristics and Biases, D. Kahneman, P. Slovic, and A. Tversky (eds.), Cambridge University Press, New York.

Waterman, D., 1985. A Guide to Expert Systems, Addison-Wesley, Reading, Mass.

Wyer, R., 1976. "An Investigation of the Relations Among Probability Estimates," Organizational Behavior and Human Performance, 15, 1-18.

Yates, J. F. and Carlson, B., 1986. "Conjunction Errors: Evidence for Multiple Judgment Procedures, Including 'Signed Summation,'" Organizational Behavior and Human Decision Processes, 37, 230-253.

Zadeh, L., 1985. "Syllogistic Reasoning as a Basis for Combination of Evidence in Expert Systems," Ninth International Joint Conference on Artificial Intelligence, 417-419.

INTERNAL DISTRIBUTION

1. L. F. Arrowood
2. C. S. Bowman
3. J. M. Bownds
4. M. A. Brown
5. D. J. Bjornstad
6. R. B. Braid
7. S. Cantor
8. R. Cantor
9. S. A. Carnes
10. K. R. Carr
11. C. V. Chester
12. J. E. Christian
13. J. L. Christian
14. T. R. Curlee
15. G. A. Dailey
16. S. J. Dale
17. L. D. Duncan
18. W. Fulkerson
19. R. T. Goeltz
20. R. K. Gryder
21. K. Hake
22. J. L. Hardee
23. M. R. Hilliard
24. E. L. Hillsman
25. E. Hirst
26. R. B. Honea
27. H. C. Hwang
28. A. F. Huntley
29. B. N. Janson
30. C. K. Johnson
31. D. W. Jones
32. J. O. Kolb
33. R. Lee
34. P. Leiby
35. G. E. Liepins
36. A. S. Loebl
37. R. S. Loffman
38. W. R. Mixon
39. V. C. Nei
40. P. J. Otaduy
41. W. J. Pennewell
42. C. H. Petrich
43. G. T. Privon
44. S. L. Purucker
45. S. Rayner
46. D. E. Reichle
47. D. T. Rizy
48. E. Rogers
49. R. M. Rush
50. M. Schoepfle
51. F. L. Sexton
52. R. B. Shelton
53. C. E. Snyder
54. E. J. Soderstrom
55. J. H. Sorensen
56. F. Southworth
57. J. P. Stovall
58. B. E. Tonn
59. A. H. Voelker
60. D. P. Vogt
61. D. Waddle
62. T. J. Wilbanks
63. A. Wolfe
64. Document Reference Section

65-66. Central Research Library

67. Laboratory Records Department
68. Laboratory Records Department (RC)
69. ORNL Patent Section

EXTERNAL DISTRIBUTION

70. J. J. Cuttica, Vice President of Research and Development, Gas Research Institute, 8600 W. Bryn Mawr Avenue, Chicago, IL 60631
71. J. P. Kalt, Professor of Economics, Kennedy School of Government, Harvard University, 79 John F. Kennedy Street, Cambridge, MA 02138
72. Barbara Kavanaugh, Center for Energy Studies, Louisiana State University, Baton Rouge, LA 70803-0301
73. Don MacGregor, MacGregor-Bates, Inc., 3820 Watkins Lane, Eugene, OR 97405

74. D. E. Morrison, Professor of Sociology, Michigan State University, 201 Berkey Hall, East Lansing, MI 48824-1111
75. R. L. Perrine, Professor, Engineering and Applied Sciences, Civil Engineering Department, Engineering I, Room 2066, University of California, Los Angeles, CA 90024
76. Cheryl Travis, Department of Psychology, University of Tennessee, Knoxville, TN 37996
77. Carl Wagner, Math Department, University of Tennessee, Knoxville, TN 37996-1300
78. Office of Assistant Manager, Energy Research and Development, DOE/ORO, Oak Ridge, TN 37830

79-80. OSTI, U. S. Department of Energy, P. O. Box 62, Oak Ridge, TN 37831

81-110. Extra copies to E. M. Schorn, 4500N, MS206

☆ U.S. GOVERNMENT PRINTING OFFICE 1988—548-118/80110