(10)

LEVEL II

TECHNICAL REPORT PTR-1042-77-7

PERCEIVED FREQUENCY OF RECURRENT EVENTS

DECISION RESEARCH • A BRANCH OF PERCEPTRONICS

Sarah Lichtenstein
Paul Slovic
Baruch Fischhoff
Mark Layman
Barbara Combs

July 1977

ADVANCED DECISION TECHNOLOGY PROGRAM

CYBERNETICS TECHNOLOGY OFFICE
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
Office of Naval Research • Engineering Psychology Programs

TECHNICAL REPORT PTR-1042-77-7

PERCEIVED FREQUENCY OF RECURRENT EVENTS

by

Sarah Lichtenstein, Paul Slovic,
Baruch Fischhoff, Mark Layman and Barbara Combs

Sponsored by

Defense Advanced Research Projects Agency
ARPA Order No. 3469

Under Subcontract from

Decisions and Designs, Incorporated

July 1977

DECISION RESEARCH
A BRANCH OF PERCEPTRONICS
1201 Oak Street
Eugene, Oregon 97401
(503) 485-2400

SUMMARY

Overview

A series of experiments involving over 650 individuals studied how people judge the frequency of death from various causes. The judgments revealed a highly consistent but systematically biased subjective scale of frequency. Possible sources of these biases as well as the implications for decision making are discussed.

Background and Approach

Decision making and planning often call for assessing the probability of future hazardous events (e.g., a swift build-up of enemy forces) or the frequency of past hazardous events (e.g., the failure rate for a particular part or command system). Considerable effort has gone into studying such judgments for relatively frequent or likely events of non-hazardous nature. Little attention has been given to the assessment of low frequencies and probabilities, such as those associated with life-threatening risks or failure rates for fail-safe systems. One reason for such inattention is the difficulty of obtaining correct answers against which judgments can be compared. The present studies look at frequency judgment in three content areas for which correct frequency tallies are available. The major focus of study is the judged frequency of death from various causes. Studies of frequency of words and occupations were also conducted for comparison purposes.

Findings

People have subjective frequency scales that are highly consistent internally, but that are systematically biased. Two

kinds of bias were identified: (a) a tendency to overestimate small frequencies and underestimate larger ones, and (b) a tendency to exaggerate the frequency of some individual events and to underestimate the frequency of others, at any given level of objective frequency. These biases were traced to a number of possible sources including disproportionate exposure, memorability or imaginability of various events. Participants in these studies were unable to correct for these sources of bias when specifically told to avoid them.

Implications

Even though people are exposed daily to information about risks and deaths, they apparently do not store this information in ways allowing them to make completely veridical estimates of frequency. Extrapolating to other areas (for which correct answers are unavailable), it is not safe to assume that experience with a particular kind of event confers the ability to make valid frequency estimates about it. The tendency to overestimate small frequencies and underestimate larger ones may be eliminated through use of a simple correction rule applied to the set of estimates. Tendencies to over- and underestimate that vary from event to event seem much more difficult to overcome, except by finding ways to help people get a better appraisal of the limits and biases of their own knowledge.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

FIGURES

TABLES

ACKNOWLEDGMENT

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by Office of Naval Research under Contracts N00014-76-C-0074 and N00014-78-C-0100 (ARPA Order Nos. 3052 and 3469) under Subcontract to Oregon Research Institute and Subcontracts 76-030-0714 and 78-072-0722 to Perceptronics, Inc. from Decisions and Designs, Inc.

1.0 INTRODUCTION

How well can people estimate the frequencies of the lethal events they may encounter in life (e.g., accidents, diseases, homicides, suicides, etc.)? More specifically, how small a difference in frequency can be reliably detected? Do people have a consistent internal scale of frequency for such events? What factors, besides actual frequency, influence people's judgments?

The answers to these questions may have great importance to society. Citizens must perceive risks accurately in order to mobilize society's resources effectively for reducing hazards and treating their victims. Official recognition of the importance of valid risk perceptions may be found in the "vital statistics" that are carefully tabulated and periodically reported to the public (see Figure 1-1). There is, however, no guarantee that these statistics are reflected in the public's intuitive perceptions.

Few studies have addressed these questions. Most investigations of perceived frequency have been laboratory experiments using sequential or simultaneous displays of lights, letters, numbers, or horizontal and vertical lines. In such tasks, people's judgments of freqency and proportion have typically been quite accurate. According to Peterson and Beach (1967), the most striking aspect of many of these studies was that the relation between estimated and actual frequency was described well by the identity function. Howell's (1973) review of the literature concluded: ". . . subjects show a remarkable facility for synthesizing and storing the repetitive attribute of event occurrences. They seem capable of maintaining a number of separate frequency streams concurrently as evidenced by the creditable accuracy of frequency retrieval" (p. 51).

Earthquakes not more numerous

GOLDEN, Colo (UPI) — The National Earthquake Information Center says recent earthquakes in populated regions may have falsely led the public to believe an abnormally high amount of quakes have been occurring.

"Recently, earthquakes have gotten more attention because they have been in populated areas," said Geophysicist Waverly Persons. "But it you look at the quakes, they all occurred in highly seismic areas where earthquakes have occurred all through histor[illegible]

Island region of the South P[illegible] C[illegible] The quake registered 6.9 [illegible]

U.S. death rate from fires sets world record

NEW YORK (UPI) — The United States is clinging to a major distinction it can hardly boast about: it has the highest death rate from fires of any free world industrialized nation.

This depress[illegible] by the Insura[illegible] in remindin[illegible] observance starting [illegible]

It's [illegible] by t[illegible] as[illegible]

Encephalitis claims lives of 24 persons

By Associated Press

Encephalitis outbreaks strikin[illegible] mainly at the elderly or the your[illegible] have claimed at least 26 lives in t[illegible] United States this year with the seas[illegible] for the disease only half over, medi[illegible] authorities say.

A type of the disease known as St. Louis strain has killed 24 pe[illegible] in Mississippi according to [illegible]

Suicides up but top killers under control

WASHINGTON (UPI) — Alth[illegible] more Americans are committing [illegible]cide today than 25 years ago, the killer diseases of 1950 have either b[illegible] eliminated or are in decline, the fede[illegible] government reported Thursday.

The report, compiled by the Natio[illegible] Center for Health Statistics and pr[illegible] sented to the Senate health subcomm[illegible] tee, said the health status of m[illegible] age[illegible] 25 to 44 has [illegible]

Botulism outbreaks linked to hard times

With millions of Americans who never before wielded a hoe now harvesting home[illegible] and canning [illegible]on, the [illegible]

This prospect and the mushrooming botulism problem has United States Departm[illegible] ture to conduct an exten[illegible] program to bring infor[illegible]

phlet, "Home and Garden [illegible]me Canning of Fruits

Fire fatalities down slightly

BOSTON (AP) — Fire deaths were down and property damage up in 1974, according to the National Fire Protection Assn.

The association says that about

Traffic deaths drop in nation

WASHINGTON (UPI) — Thanks in part to the energy problem traffic deaths on the nation's roads and highways decreased by 9,550 last year, the government reports.

The National Highway Traffic Safe[illegible]

Bubonic plague expected to set record for year

FORT COLLINS, Colo. (UPI) — [illegible]bonic plague is the "Black Death" killed millions in Europe in the Mi[illegible] Ages. Now, U.S. health officials [illegible] there are indications this will be [illegible] ord year for the disease in the U[illegible] States [illegible] deaths fro[illegible]

FIGURE 1-1. EAT, DRINK AND BE MERRY

Similarly, Estes (1976) observed that subjects in probability-learning experiments were "extremely efficient" (p. 51) at acquiring relative-frequency information.

Despite these optimistic conclusions, some studies have found inaccuracies. For example, Attneave (1953) and Hintzman (1969) found that judged frequency increased with the log of the true frequency. Still other studies have suggested some cognitive processes that could lead to even more serious errors in judgments of lethal events. In this regard, Postman (1964) noted that frequency learning is typically incidental learning, which is strongly influenced by selective attention. Estes (1976) observed that accurate learning of frequencies requires the learner to "attend to and encode occurrences of all the alternative events with equal uniformity or efficiency" (p. 53). Underwood (1969) found that items were judged more frequent under conditions of distributed rather than massed practice and Hintzman (in press) discussed a great deal of evidence showing that apparent frequency of an item increases with greater spacing between its repetitions in a list. Any of these factors could bias judgments about the frequencies of causes of death. Events that capture our attention and "stick in our mind," like homicide, may appear more frequent than they are. Rare events may be overestimated because their appearances are well spread and distinct. Catastrophic (multi-fatality) events might be overestimated because of their salience or underestimated because of massed presentation.

Tversky and Kahneman (1973) have argued that people judge the probability or frequency of an event by the ease with which relevant instances can be retrieved from memory or imagined. Reliance on memorability and imaginability as a cue for frequency was called the "availability" heuristic. In the context of lethal events, "availability" implies that direct

experience with an event will affect one's judgments of its frequency, as will indirect exposure to the event via movies, books, television, newspapers, etc. Thus we might expect that the frequencies of dramatic events such as cancer, homicide or multiple-death catastrophes, which tend to be publicized disproportionately, would be overestimated, while the frequencies of less dramatic killers would be underestimated.

In summary, experimental research shows that although people are very good at tracking event frequencies, the potential exists for serious misjudgment. Even without the ambiguity of this conclusion, the implications of these studies for judgments regarding causes of death would be unclear. Lethal events are emotion-laden stimuli experienced in many different contexts, over the course of a lifetime. Some of these events occur thousands of times more frequently than others. No laboratory experiments have even approximated these conditions.

Perhaps more relevant are field surveys by several geographers (Burton, Kates & White, in press; Kates, 1962, in press; White, 1974). These studies have indicated (a) that people misperceive the hazards posed by floods, earthquakes, hurricanes and drought; (b) that more frequent hazards are perceived more accurately; and (c) accuracy is increased by both the recency of the hazard's last major occurrence and its impact on one's livelihood.

Judgments concerning the probabilities and frequencies of real-life events have also been studied by Selvidge (Note 1). In one phase of her research, five subjects first ranked several sets of accidents and crimes according to frequency,

and then estimated the absolute frequencies. Although her subjects were fairly good at ordering the events, they did a poor job of assigning absolute frequencies. She also found a great amount of variability across subjects, event categories, and response modes. This variability and her small sample size led Selvidge to advocate that these issues be investigated on a much larger scale. The present study does this.

Five experiments are reported here. The first two examine the accuracy of comparative judgments, using a paired-comparison format. The third evaluates judgments of absolute frequency. The fourth examines the role that several aspects of availability may play in determining such judgments. The fifth explores the degree to which subjects can overcome their errors when informed of the nature of their biases.

2.0 EXPERIMENT 1: PAIRED COMPARISON JUDGMENTS OF LETHAL EVENTS

The first experiment investigated the accuracy of relative-frequency judgments for various causes of death.

2.1 Method

2.1.1 Stimuli. Table 2-1 shows the stimulus events, 41 causes of death, and gives, for each item, the frequency of death per 10^8 United States residents per year, based on reports prepared by the National Center for Health Statistics for the years 1968-1973.[1] These events were chosen to represent the range of frequencies of causes of death for which yearly statistics are available. Obscure or unfamiliar causes were excluded, as were causes showing large fluctuations from year to year. For the few chosen events that showed a systematic trend across years (e.g., homicide, which increased from 7300 per 10^8 in 1968 to 9400 per 10^8 in 1973), the average over the last two years was used.

From these 41 causes of death, 106 pairs were constructed such that (a) each cause appeared in approximately six pairs and (b) the ratios of relative frequencies (comparing the more to the less frequent cause of death) varied systematically from 1.25 : 1 (example: fireworks vs. measles) to about 100,000 : 1 (example: stroke vs. botulism). Five pairs included smallpox as the less frequent cause of death. Since no one in the United States has died of smallpox since 1949, the rate shown in Table 2-1 is zero, and no ratio comparing any other disease with smallpox can be defined.

---

[1] For convenience, these frequencies are referred to in this paper as "the true frequencies," although we recognize that they are statistical estimates.

TABLE 2-1

CAUSES OF DEATH MASTER LIST

	Rate/10^8		Rate/10^8
Smallpox	0	Firearm accident	1,100
Poisoning by vitamins	.5	Poisoning by solid or liquid	1,250
Botulism	1	Tuberculosis	1,800
Measles	2.4	Fire & flames	3,600
Fireworks	3	Drowning	3,600
Smallpox vaccination	4	Leukemia	7,100
Whooping cough	7.2	Accidental falls	8,500
Polio	8.3	Homicide	9,200
Venomous bite or sting	23.5	Emphysema	10,600
Tornado	44	Suicide	12,000
Lightning	52	Breast cancer	15,200
Non-venomous animal	63	Diabetes	19,000
Flood	100	Motor vehicles (car, truck or bus) accident	27,000
Excess cold	163	Lung cancer	37,000
Syphilis	200	Cancer of the digestive system	46,600
Pregnancy, childbirth, and abortion	220	All accident	55,000
Infectious hepatitis	330	Stroke	102,000
Appendicitis	440	All cancer	160,000
Electrocution	500	Heart disease	360,000
Motor vehicle-train collision	740	All disease	849,000
Asthma	920		

In the results that follow, all analyses employing ratios of true frequencies (called "true ratios") exclude the five pairs involving smallpox.

2.1.2 Subjects. Two groups of subjects participated. The first, hereafter referred to as the "college students," consisted of 51 males and 60 females who answered an ad in the University of Oregon campus newspaper. The second consisted of 77 female members of the Eugene, Oregon, Chapter of the League of Women Voters, a group representative of the best-informed citizens in the community. All subjects were paid for participating. The data were collected from the students in the autumn of 1974 and from the League members in the spring of 1975.

The order of the 106 pairs and of the two causes within each pair was determined randomly. All subjects saw the same random order.

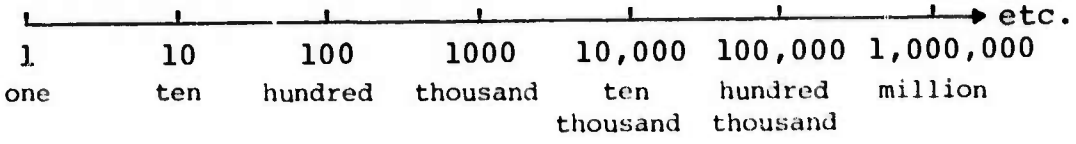
2.1.3 Instructions. The subjects' instructions read as follows:

> Each item in part one consists of two different possible causes of death. The question you are to answer is: Which cause of death is more likely? We do not mean more likely for you, we mean more likely in general, in the United States.
>
> Consider all the people now living in the United States--children, adults, everyone. Now supposing we randomly picked just one of those people. Will that person more likely die next year from cause A or cause B? For example: Dying in a bicycle accident versus dying from an overdose of heroin. Death from each cause is remotely possible. Our question is, which of these two is the more likely cause of death?

For each pair of possible causes of death, A and B, we want you to mark on your answer sheet which cause you think is MORE LIKELY.

Next, we want you to decide how many times more likely this cause of death is, as compared with the other cause of death given in the same item. The pairs we use vary widely in their relative likelihood. For one pair, you may think that the two causes are equally likely. If so, you should write the number 1 in the space provided for that pair. Or, you may think that one cause of death is 10 times, or 100 times, or even a million times as likely as the other cause of death. You have to decide: How many <u>times</u> as likely is the more likely cause of death? Write the number in the space provided. If you think it's twice as likely, write 2. If it's 10 thousand times as likely, write 10,000, and so forth.

At the top of the answer sheet, we have drawn a little scale that looks like this:



The scale is there to give you an idea of the kinds of numbers you might want to use. You don't have to use exactly those numbers. You could write 75 if you think that the more likely cause of death is 75 times more likely than the other cause, or 500, if you think that the more likely cause of death is 500 times more likely than the other.

For some pairs, you may believe that one cause of death is just a little bit more likely than the other cause of death. For this situation, you will have to use a decimal point in your answer:

1.1 means that the more likely cause is 10% more likely than the other cause.

1.2 means 20% more likely.

1.5 means 50% more likely, or half again as likely.

⋮

1.8 means 80% more likely.

⋮

2 means twice as likely, which is the same as 100% more likely.

⋮

2.5 means two and a half times as likely.

In addition, the following glossary was provided to insure that the subjects understood what was included in some possibly ambiguous categories:

All accidents: includes any kind of accidental event; excludes diseases and natural disasters (floods, tornadoes, etc.).

All cancer: includes leukemia.

Cancer of the digestive system: includes cancer of stomach, alimentary tract, esophagus and intestines.

Excess cold: freezing to death or death by exposure.

Non-venomous animal: dogs, bears, etc.

Venomous bite or sting: caused by snakes, bees, wasps, etc.

2.2 Results

2.2.1 Accuracy. Two measures were computed for each pair of causes of death, the percentage of subjects who correctly selected the more likely item and the geometric mean of the subjects' ratio judgments. For any subject who did not correctly select the more likely cause of death, the

inverse of the judged ratio was used in calculating the geometric mean. For example, death by fireworks is more frequent than death from measles. If a subject said measles was 5 times more likely to cause death than fireworks, the inverse, .2, was used. The two summary measures, percentage correct and the geometric mean of the ratio judgments, are shown for all 106 pairs for both groups of subjects in Table 2-2.

Examination of Table 2-2 illustrates the many, often severe, misconceptions held by both the college students and the League members. For example, even though stroke causes 85% more deaths than all accidents combined (pair 37, true ratio = 1.85), only 20% of the students and 23% of the League members judged stroke to be more likely. The geometric mean of the ratio judgments was only .04 for the students, indicating that, on the average, they believed that accidents were 25 times (1 ÷ .04) more frequent. Tornadoes were seen by the student subjects as more frequent killers than asthma, even though the latter is 21 times more likely (pair 61). Death by lightning was perceived as less likely than by botulism even though it is 52 times more frequent (pair 71). Death by asthma was judged only slightly more frequent than death by botulism (pair 91), even though it is over 900 times more frequent! Accidental deaths were perceived by the students to be about as likely as death from disease despite a true ratio of 15.4 for diseases over accidents (pair 69).

Some errors were in the opposite direction: A large percentage of subjects knew which cause of death was more likely, but the ratios given were far too large. For

TABLE 2-2

RESULTS OF PAIRED-COMPARISON JUDGMENTS FOR CAUSES OF DEATH

Pair #	Less Likely	More Likely	Rate of Less Likely	True Ratio	% Correct		Geometric Mean	
					Students	LWV	Students	LWV
1	Smallpox	Botulism	0	-	77	88	5.54	18.3
2	Smallpox	Measles	0	-	60	86	1.52	16.0
3	Smallpox	Smallpox Vaccination	0	-	26	56	.08	1.08
4	Smallpox	Polio	0	-	64	74	2.76	5.90
5	Smallpox	Heart Disease	0	-	97	97	1130.0	13263
6	Measles	Fireworks	2.4	1.25	28	35	.27	.44
7	Fireworks	Smallpox Vaccination	3	1.33	36	29	.47	.29
8	Lightning	Non-venomous Animal	52	1.21	55	58	.71	1.89
9	Excess Cold	Syphilis	163	1.23	72	81	5.93	23.7
10	Asthma	Firearm Accident	920	1.20	60	78	11.0	14.1
11	Leukemia	Accidental Falls	7100	1.20	46	64	.78	2.45
12	Accidental Falls	Emphysema	8500	1.25	41	39	.65	.45
13	Homicide	Suicide	9200	1.30	32	30	.19	.14
14	Breast Cancer	Diabetes	15200	1.25	23	27	.13	.14
15	Lung Cancer	Stomach Cancer	37000	1.25	25	44	.31	.81
16	Stomach Cancer	All Accidents	46600	1.19	96	87	28.1	29.4
17	Tornado	Non-venomous Animal	44	1.43	20	21	.10	.15
18	Flood	Excess Cold	100	1.63	37	42	.51	.34
19	Hepatitis	Electrocution	330	1.52	45	38	.61	.44
20	Electrocution	Motor-Train Collision	500	1.48	45	52	1.16	2.40
21	Poisoning	Tuberculosis	1250	1.44	26	12	.19	.03
22	Leukemia	Emphysema	7100	1.49	47	47	.58	1.02
23	Accidental Falls	Suicide	8500	1.41	58	51	1.58	.64
24	Suicide	Diabetes	12000	1.58	25	11	.09	.20
25	Diabetes	Motor Vehicle Accident	19000	1.42	99	96	356.0	99.6
26	Lung Cancer	All Accidents	37000	1.49	81	90	11.1	34.7
27	Stroke	All Cancer	102000	1.57	83	75	21.0	8.20
28	Poison by Vitamin	Botulism	.5	2.00	68	82	4.08	14.0
29	Lightning	Flood	52	1.92	85	87	18.6	14.1
30	Flood	Syphilis	100	2.00	57	73	1.74	6.60

TABLE 2-2 (CONT'D.-2)

Pair #	Less Likely	More Likely	Rate of Less Likely	True Ratio	% Correct		Geometric Mean	
					Students	LWV	Students	LWV
31	Pregnancy, etc.	Appendicitis	220	2.00	17	10	.10	.07
32	Appendicitis	Asthma	440	2.09	50	71	1.00	6.65
33	Tuberculosis	Fire & Flames	1800	2.00	81	92	10.5	38.9
34	Tuberculosis	Drowning	1800	2.00	67	83	2.98	19.3
35	Leukemia	Breast Cancer	7100	2.14	58	60	1.48	2.98
36	Breast Cancer	Lung Cancer	15200	2.43	82	71	6.42	2.74
37	All Accidents	Stroke	55000	1.85	20	23	.04	.13
38	All Cancer	Heart Disease	160000	2.25	55	68	.89	1.88
39	Poison by Vitamin	Measles	.5	4.80	66	74	2.50	5.45
40	Polio	Tornado	8.3	5.30	71	86	4.26	18.0
41	Non-venomous Animal	Hepatitis	63	5.24	84	90	9.76	15.1
42	Syphilis	Firearm Accident	200	5.50	66	77	4.38	16.6
43	Pregnancy, etc.	Firearm Accident	220	5.00	62	68	2.14	6.66
44	Motor-Train Collision	Fire & Flames	740	4.86	67	90	2.45	14.2
45	Motor-Train Collision	Drowning	740	4.86	67	77	1.97	7.17
46	Tuberculosis	Homicide	1800	5.11	91	91	25.2	72.9
47	Emphysema	All Accidents	10600	5.19	88	95	269	107
48	Diabetes	Stroke	19000	5.37	91	84	29.3	46.6
49	Measles	Venomous Bite/Sting	2.4	9.79	65	68	1.68	3.11
50	Smallpox Vaccination	Tornado	4	11.00	84	91	16.8	22.7
51	Venomous Bite/Sting	Pregnancy, etc.	23.5	9.36	77	82	7.27	11.6
52	Lightning	Electrocution	52	9.62	88	92	23.1	26.6
53	Hepatitis	Fire & Flames	330	10.90	67	78	4.22	11.2
54	Hepatitis	Drowning	330	10.90	68	71	4.55	7.80
55	Poisoning	Suicide	1250	9.60	70	73	5.50	2.59
56	Tuberculosis	Diabetes	1800	10.56	53	90	1.47	14.7
57	Emphysema	Stroke	10600	9.62	81	74	10.5	10.2
58	Lung Cancer	Heart Disease	37000	9.73	81	95	3.71	24.4
59	Measles	Tornado	2.4	18.3	68	70	5.63	4.67
60	Polio	Excess Cold	8.3	19.6	63	79	1.64	6.05

TABLE 2-2 (CONT'D.-3)

Pair #	Less Likely	More Likely	Rate of Less Likely	True Ratio	% Correct		Geometric Mean	
					Students	LWV	Students	LWV
61	Tornado	Asthma	44	20.9	42	68	.36	3.53
62	Non-venomous Animal	Poisoning	63	19.8	95	95	17.1	64.9
63	Hepatitis	Leukemia	330	21.5	75	79	12.5	14.7
64	Appendicitis	Homicide	440	20.9	91	97	72.7	105
65	Motor-Train Collision	Breast Cancer	740	20.5	70	83	4.90	20.9
66	Poisoning	Motor Vehicle Accident	1250	21.6	95	94	388	304
67	Tuberculosis	Lung Cancer	1800	20.6	85	99	23.3	145
68	Diabetes	Heart Disease	19000	18.9	97	97	127	206
69	All Accidents	All Diseases	55000	15.4	57	79	1.62	11.6
70	Poison by Vitamin	Venomous Bite/Sting	.5	47.0	75	79	4.04	16.5
71	Botulism	Lightning	1	52.0	37	45	.30	.32
72	Whooping Cough	Hepatitis	7.2	45.8	88	91	7.66	12.4
73	Venomous Bite/Sting	Firearm Accident	23.5	46.8	85	87	9.94	35.7
74	Venomous Bite/Sting	Poisoning	23.5	53.2	78	83	5.21	11.3
75	Syphilis	Homicide	200	46.0	86	82	31.7	44.2
76	Pregnancy, etc.	Suicide	220	54.5	61	79	14.8	12.7
77	Electrocution	Motor Vehicle Accident	500	54.0	97	99	539	909
78	Asthma	Stomach Cancer	920	50.4	86	90	36.4	43.4
79	Leukemia	Heart Disease	7100	50.7	90	99	33.8	184
80	Poison by Vitamin	Lightning	.5	104	50	61	1.45	2.00
81	Measles	Pregnancy, etc.	2.4	92	84	79	13.4	9.44
82	Whooping Cough	Motor-Train Collision	7.2	103	85	95	15.2	70.1
83	Flood	Homicide	100	92	91	94	81.7	294
84	Excess Cold	Breast Cancer	163	93	80	95	26.7	255
85	Syphilis	Diabetes	200	95	64	71	2.36	4.52
86	Appendicitis	Stomach Cancer	440	105	89	95	30.2	83.7
87	Drowning	Heart Disease	3600	100	92	95	56.9	272
88	Fire & Flames	Heart Disease	3600	100	93	94	79.7	252
89	Accidental Falls	All Disease	8500	100	91	90	324	614
90	Poison by Vitamin	Electrocution	.5	1000	85	86	30.7	27.7

TABLE 2-2 (CONT'D.-4)

Pair #	Less Likely	More Likely	Rate of Less Likely	True Ratio	% Correct		Geometric Mean	
					Students	LWV	Students	LWV
91	Botulism	Asthma	1	920	59	75	1.50	9.29
92	Botulism	Firearm Accident	1	1100	88	94	29.9	189
93	Whooping Cough	Leukemia	7.2	986	94	99	45.2	166
94	Polio	Accidental Falls	8.3	1024	93	95	17.5	198
95	Flood	Stroke	100	1020	89	90	43.4	186
96	Excess Cold	All Cancer	163	982	95	99	1490	4337
97	Botulism	Emphysema	1	10600	86	94	24	190
98	Measles	Motor Vehicle Accident	2.4	11250	96	99	1070	2765
99	Fireworks	Motor Vehicle Accident	3	9000	95	97	1430	5268
100	Polio	Stroke	8.3	12289	95	99	164	2265
101	Tornado	Heart Disease	44	8182	95	97	348	2396
102	Non-venomous Animal	All Disease	63	13476	98	100	5600	33521
103	Poison by Vitamin	Stomach Cancer	.5	92800	92	99	103	578
104	Botulism	Stroke	1	102000	96	95	106	468
105	Fireworks	Heart Disease	3	120000	96	99	1530	5779
106	Smallpox Vaccination	Heart Disease	4	90000	99	100	3610	12244

example, death by a motor vehicle accident is only 1.4 times more likely than death from diabetes (pair 25), not 356 times more likely (the students' geometric mean) or 100 times more likely (League members).

Subjects' overall level of performance was not quite as bad as these examples suggest. They were generally able to identify the more frequent cause of death when the true ratio was 2 : 1 or greater. Below 2 : 1, however, discrimination was often poor, as shown in Figures 2-1 and 2-2, which compare the percentage of correct discriminations with the log true ratio for the two groups of subjects (101 pairs, excluding smallpox).

Accuracy as measured by percentage correct was slightly higher for events higher in statistical frequency. The partial correlation between percentage correct and log frequency of the less likely event, holding true ratio constant, was .24 ($\underline{z}$ = 2.48; 1 tailed $p < .01$) for the college students, and .19 ($\underline{z}$ = 1.62; 1 tailed $p < .06$) for the League members.

The geometric means of the likelihood judgments were only moderately related to the true ratios of frequencies, as shown in Figures 2-3 and 2-4 (101 pairs, excluding smallpox). For example, the college students produced mean ratios in the range of 100 : 1 to 500 : 1 for pairs with true ratios as small as 1.5 : 1 and as large as 100,000 : 1! Conversely, pairs having true ratios of about 2 : 1 had geometric mean judgments ranging from 25 : 1 in the wrong direction to over 300 : 1 in the right direction! The geometric means were somewhat more accurate for the League

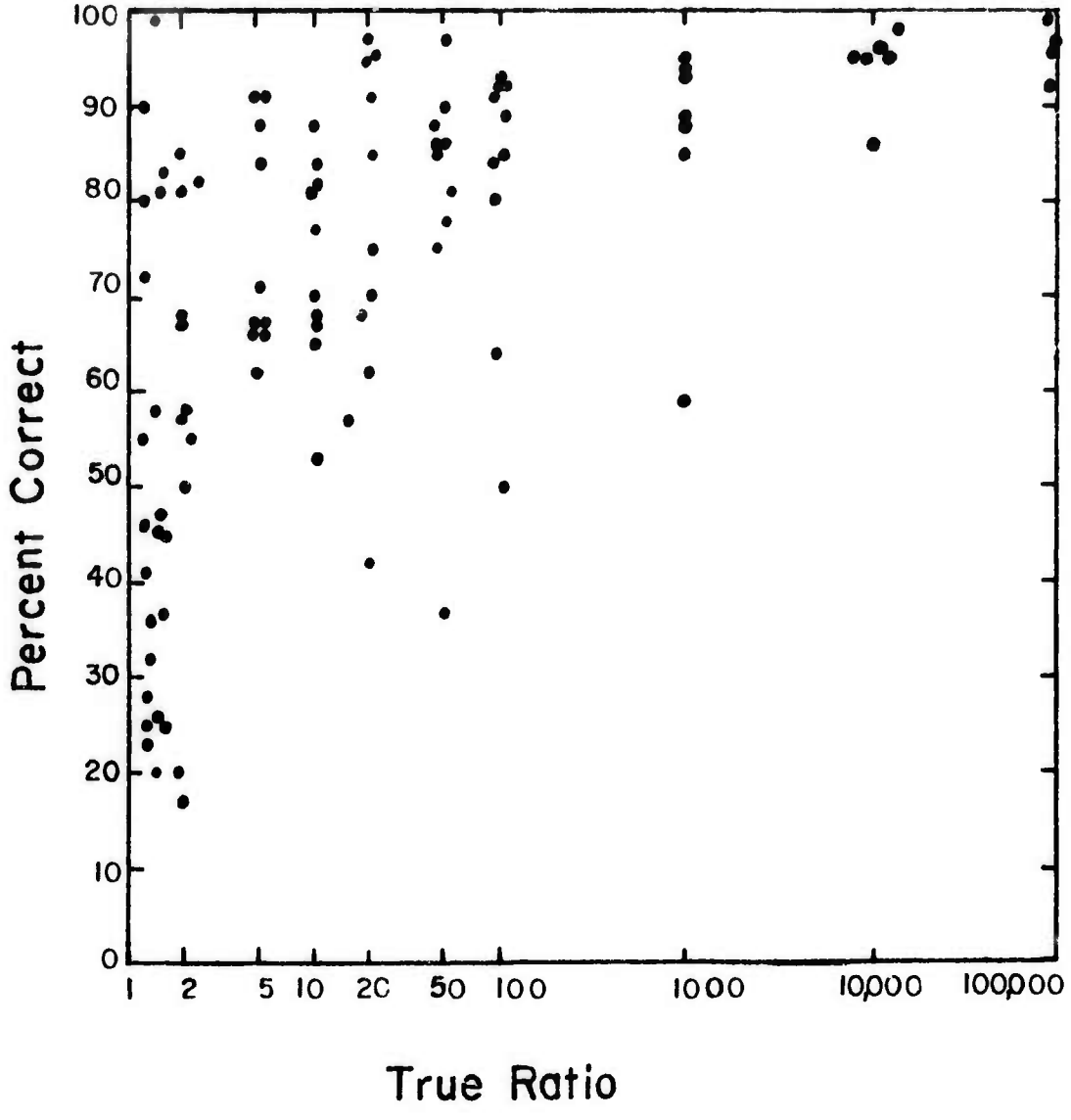


FIGURE 2-1. PERCENT CORRECT AS A FUNCTION OF TRUE RATIO FOR 101 PAIRED CAUSES OF DEATH: STUDENTS

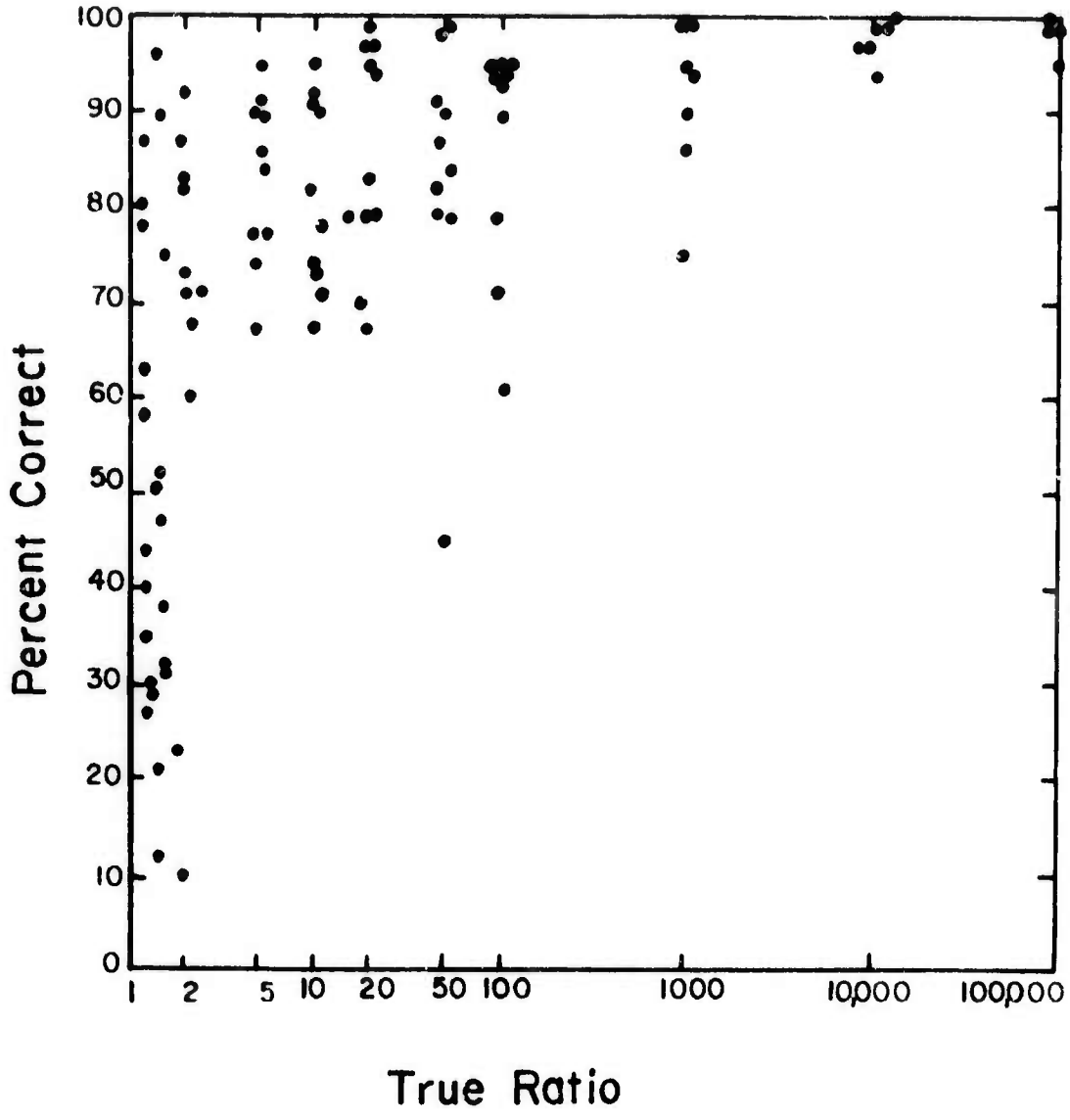


FIGURE 2-2. PERCENT CORRECT AS A FUNCTION OF TRUE RATIO FOR 101 PAIRED CAUSES OF DEATH: LEAGUE OF WOMEN VOTERS

members, but still were far from optimal. The correlation between log geometric mean judged ratio and log true ratio was .69 for the students and .75 for the League members. The regression lines (shown as dashed lines in Figures 2-3 and 2-4) were both too flat.

2.2.2 Secondary bias. The regression lines shown in Figures 2-3 and 2-4 capture what we will call "primary bias": a tendency to underestimate large ratios. In addition, the data showed a "secondary bias": different pairs with the same true ratio had quite different judged ratios. One measure of this secondary bias is the signed difference between the log geometric mean for a pair and its log geometric mean as predicted by the regression equation. (This measure is equivalent to the vertical distance between a point in Figure 2-3 or 2-4 and the dashed regression line.) A positive value indicates that the ratio judgments for that pair were large relative to the general relationship between the judged ratio and the true ratio. A negative value indicates relative underestimation or estimation in the wrong direction. As measured by these residual values, secondary bias was highly consistent across the two groups of subjects: the between-group correlation of the residuals was .90 (over 101 pairs). Further analysis of secondary bias will be presented later in the paper.

2.2.3 Consistency. Even though they were often inaccurate, subjects' mean responses revealed a consistent subjective ordering for the causes of death. There were 18 triads (involving 29 of the 41 causes of death) of the form {A vs. B, B vs. C, A vs. C} within the 106 pairs (for example, All Accidents paired with Stroke, Stroke paired with

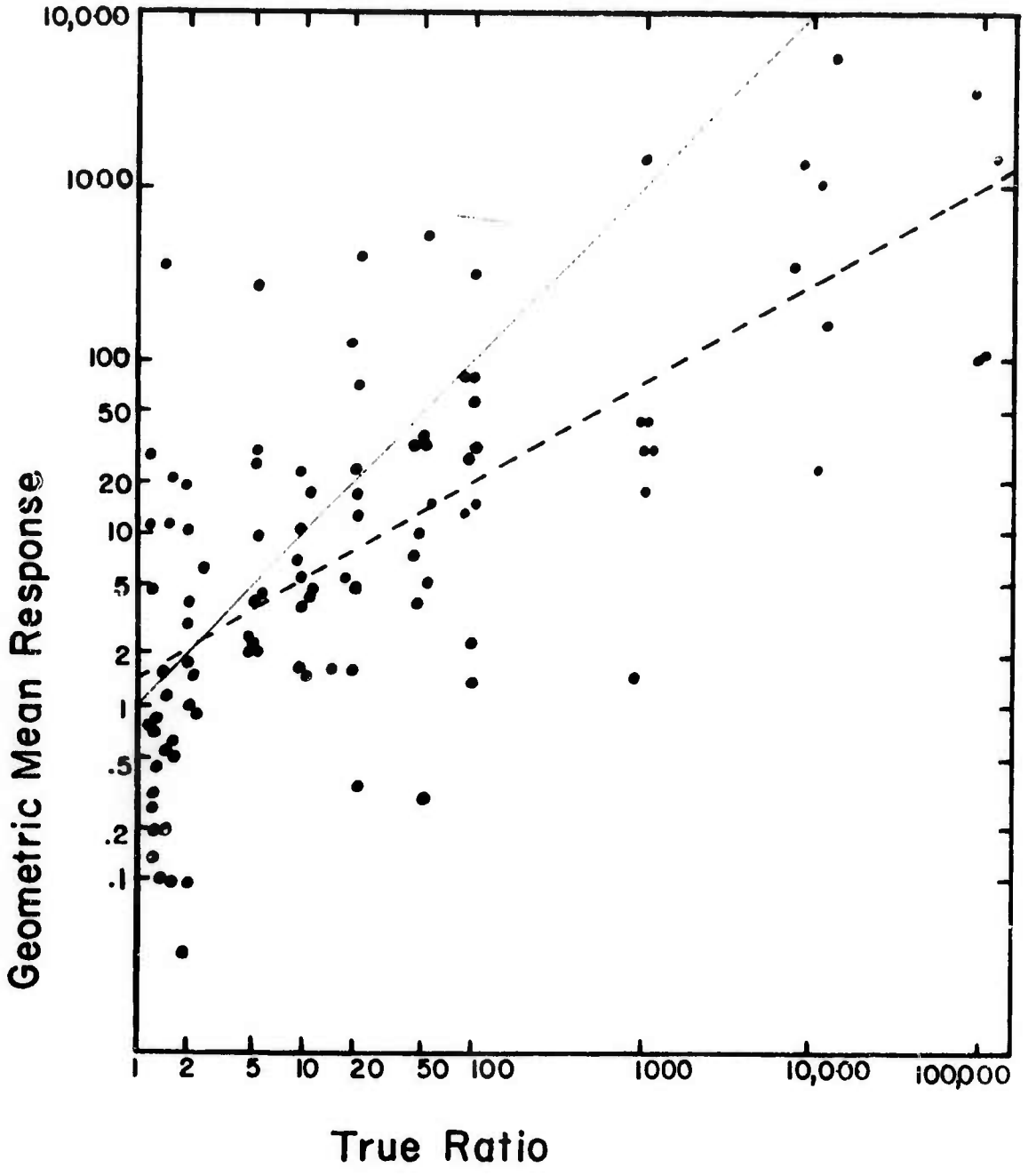


FIGURE 2-3. GEOMETRIC MEAN RATIO AS A FUNCTION OF TRUE RATIO FOR 101 PAIRED CAUSES OF DEATH: STUDENTS
DASHED LINE IS THE REGRESSION LINE

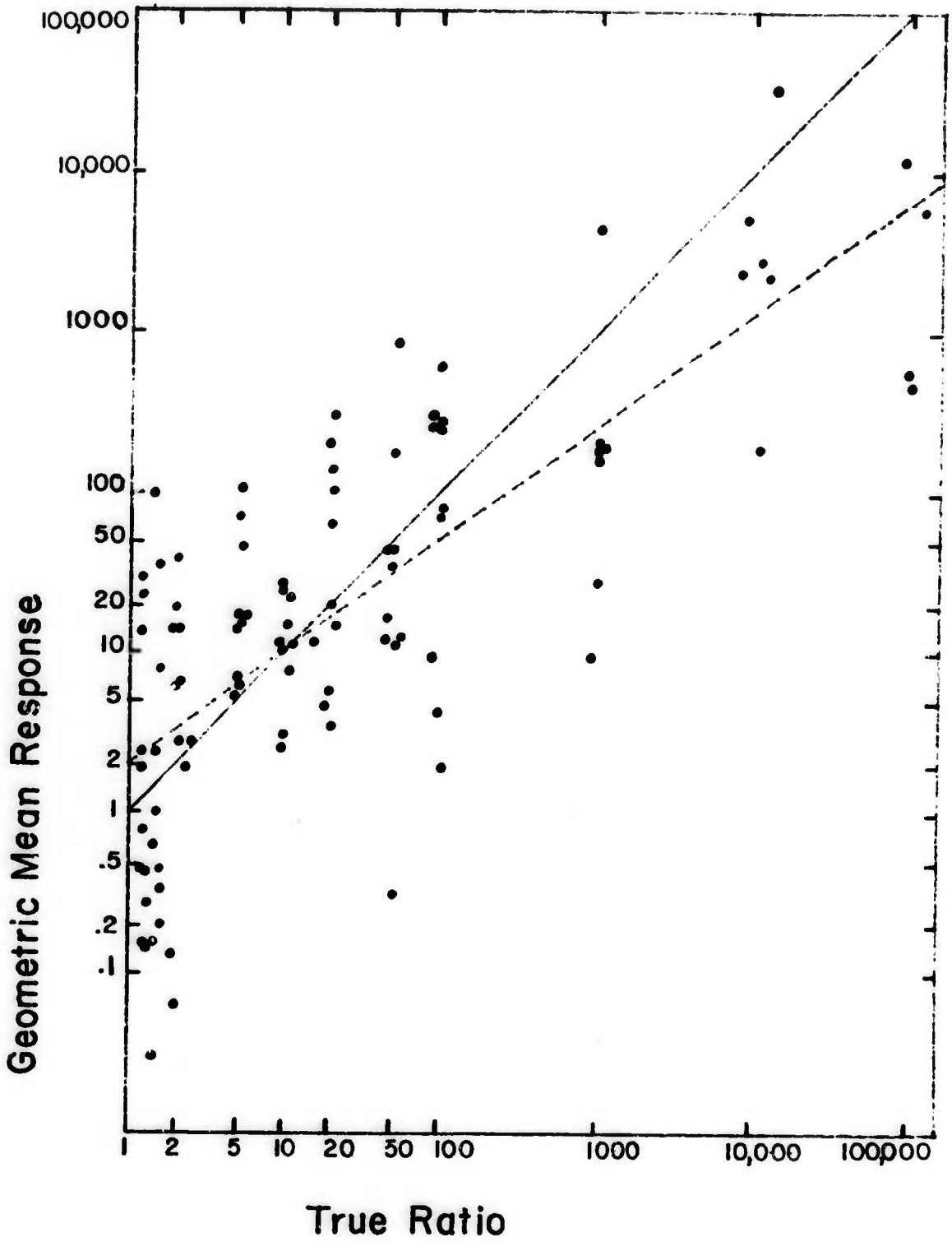


FIGURE 2-4. GEOMETRIC MEAN RATIO AS A FUNCTION OF TRUE RATIO FOR 101 PAIRED CAUSES OF DEATH: LEAGUE OF WOMEN VOTERS DASHED LINE IS THE REGRESSION LINE

Emphysema, and Emphysema paired with All Accidents). For such triads, we asked, "Were the choice percentages transitive?" and "Were the geometric means consistent?" The answer to both these questions was "yes" for the triads described above. The data were as follows:

A Majority of Student Subjects Said:	Choice Percentage	Geometric Mean Likelihood Ratio
All Accidents more likely than Stroke	60	26.3
Stroke more likely than Emphysema	81	10.5
All Accidents more likely than Emphysema	88	269.0

This triad exhibits strong stochastic transitivity[2]: The percentage of subjects judging All Accidents to be more likely than Emphysema was 88%, greater than either of the other two percentages. The consistency of the geometric means is shown by the similarity of the third mean (269) to the product of the first two means (276). Thus, the group showed a clear subjective ordering: Emphysema < Stroke < All Accidents. The true order, however, is Emphysema < All Accidents < Stroke. These results are typical of all 36 triads analyzed (18 triads each for college students and League members). The choice percentages exhibited weak stochastic transitivity for every triad; strong stochastic transitivity was satisfied for 27 out of 36 triads.

---

[2] Three levels of stochastic transitivity may be distinguished (cf. Coombs, Dawes & Tversky, 1970, p. 156). For any three stimuli, x, y and z, assume that $p(x,y) \geq \frac{1}{2}$ (i.e., that the proportion choosing x over y is greater than or equal to .5) and that $p(y,z) \geq \frac{1}{2}$. Then strong stochastic transitivity requires that $p(x,z) \geq \max [p(x,y), p(y,z)]$, moderate stochastic transitivity requires that $p(x,z) \geq \min [p(x,y), p(y,z)]$, while weak stochastic transitivity requires only that $p(x,z) \geq \frac{1}{2}$.

The consistency of the ratio judgments was measured by comparing the log of the geometric mean ratio for pair A : C in each triad with the log of the product of the geometric mean ratios for A : B and B : C. The relationship was linear with r = .99 (slope = 1.10; intercept = .83) for the college students and r = .97 (slope = 1.05; intercept = 1.09) for the League members.[3] These results suggest that as a group, these subjects exhibited an interval scale of subjective frequency.

2.2.4 Between-group comparisons. The responses of the students and the League members were highly similar. Across all 106 pairs, the correlation between the two groups was .93 for both percentage correct and geometric mean judged ratio. The high correlation between the two groups' secondary bias residuals is further evidence of this similarity. The League members had a somewhat higher percentage correct than the students (mean 76.8 vs. 71.3); their percentage correct was higher for 80 pairs, equal for 5 pairs, and lower for 21 pairs (sign test; $p < .001$). For the ratio judgments, however, the League members did not perform significantly better than the students; the geometric mean of their ratio judgments was closer to the true ratio for only 62 of the 106 pairs (sign test; $z = 1.65$, $p > .10$).

2.2.5 Individual differences. Table 2-3 shows the variability of individual subjects' performance. The first two rows indicate the slight superiority of the League members with respect to percentage correct. Note that no subject did worse than chance (50%) nor better than 90% correct. The

---

[3] These correlations were calculated on log data. However, for ease of interpretation, the intercepts given here are the antilogs.

TABLE 2-3

INDIVIDUAL DIFFERENCES FOR THE PAIRED-COMPARISON TASK

Measure	Group	Frequency Distribution Across Subjects				
		Lowest	1st Quartile	Median	3rd Quartile	Highest
% Correct	Students	55.7	67.9	70.8	75.4	83.9
	LWV	60.4	73.6	77.4	80.7	88.7
Correlation of log judged ratio with log true ratio	Students	-.11	.34	.45	.51	.72
	LWV	.10	.41	.51	.58	.80
Geometric mean error ratio	Students	7.48	15.0	22.5	45.4	556
	LWV	5.43	13.2	17.6	37.4	2693
No. of transitive triads (out of 18)	Students	12	16	17	18	18
	LWV	14	16	17	18	18

next two rows, which give the correlations between log judged ratio and log true ratio over 101 items, indicate that few subjects showed any appreciable ability to perform the ratio estimation task.

The next two rows give the geometric means of the error ratios for individual subjects. An error ratio is the ratio of the judgment to the truth, or vice versa, whichever is greater than 1. A subject who always gave a judged ratio off by a factor of 10, i.e., either 10 times as large or a tenth as large as the true ratio, would have a mean error ratio of 10. The median student subject erred by a factor of 22.5, while the median League member erred, on the average, by a factor of 17.6.

The next section of Table 2-3 shows the number of transitive triads (out of 18) for each subject. Only about one subject in four in each group had more than one intransitivity. Thus, the strong internal consistency found in the group data is repeated in the individual data.

3.0 EXPERIMENT 2: PAIRED COMPARISON JUDGMENTS OF WORDS AND OCCUPATIONS

In order to test whether the primary results of Experiment 1 were unique to the set of stimuli used, Experiment 1 was repeated using pairs of words and pairs of occupations as stimuli.

3.1 Method

3.1.1 Stimuli. The list of words studied is shown in Table 3-1, along with their frequency of occurrence per 10^6 words of English text. These frequencies represent an average from two separate sources. One source, the Lorge magazine count (Thorndike & Lorge, 1944), analyzed frequencies from a sample of about a million words from each of five major magazines between the years 1927 and 1938. The second source (Kučera & Francis, 1967) analyzed 500 samples of about 200 words each, taken from a wide variety of materials, ranging from newspapers to scientific journals and from popular romantic fiction to abstruse philosophical discussions. For the words in Table 3-1, the frequencies estimated by the two sources agreed closely. From this list, 100 pairs of words were selected, with true ratios ranging from 1.19 ("of" vs. "to") to 6126 ("the" vs. "cork").

The list of occupations studied is shown in Table 3-2, along with their frequency of occurrence among 10^8 employed U.S. civilian citizens. These frequencies were derived from a report compiled by the U.S. Bureau of the Census (1972). From the list, 95 pairs were selected, with true ratios

TABLE 3-1

WORDS MASTER LIST

Word	Rate/10^6
The	61,260
Of	34,716
And	29,834
To	25,892
In	19,032
That	11,483
He	10,246
For	9,118
With	7,300
On	6,730
From	4,044
When	2,807
Out	2,565
Time	1,751
Two	1,368
After	1,152
People	821
Again	730
Once	578
Next	455
Half	358
Result	222
Music	182
Couple	125
Hit	104
Proud	70
Dull	46
Tent	26
Cork	10
Jug	7
Bun	2

TABLE 3-2

OCCUPATIONS MASTER LIST

Occupation	Rate/10^8
Secretary	3,529,680
Elementary or Secondary School Teacher	3,155,206
Retail Sales Clerk	2,967,880
Truck Driver	1,802,169
Waiter or Waitress	1,331,616
Registered Nurse	1,083,800
Auto Mechanic	1,051,250
College or University Teacher	635,138
Electrician	611,935
Telephone Operator	531,655
Physician	436,322
Lawyer	339,829
Letter Carrier	329,866
Bus Driver	308,205
Bartender	246,584
Computer Programmer	210,750
Librarian	159,172
Baker	142,634
Bulldozer Operator	115,537
Garbage Collector	93,290
Upholsterer	81,118
Architect	73,418
Dietitian	52,422
Airline Purser, Steward, or Stewardess	43,891
Air Traffic Controller	33,040
Airline Pilot or Copilot	32,787
Psychiatrist	28,191
Veterinarian	25,387
Motion Picture Projectionist	20,198
Judge	16,001
FBI Special Agent	10,320
Rabbi	8,491
Embalmer	6,203
EEG Technician	3,919
Jockey	2,065
Nuclear Reactor Operator	1,568
Lay Midwife	882

ranging from 1.15 (garbage collector vs. upholsterer) to 1229 (registered nurse vs. lay midwife).

3.1.2 Subjects and instructions. The subjects were college students recruited via a campus newspaper advertisement and paid for their participation. One hundred eleven subjects judged the word pairs, and a different group of 118 individuals judged occupations. The instructions for words and occupations paralleled those for causes of death. For pairs of words, the subjects were asked to judge which word is more likely to be sampled at random from common writing (magazines and books, fiction, nonfiction, scientific, nonscientific, etc.) in the United States, and to indicate how many times more likely the more frequent word is than the other word in the pair. For occupations, subjects were asked to indicate whether an employed U.S. citizen picked at random is more likely to be working as an A or a B, and how many times more likely the more frequent occupation is than the other occupation in the pair.

3.2 Results

3.2.1 Accuracy. Figures 3-1 and 3-2 show the relationship between percentage correct and true ratio, while geometric mean ratio judgments are plotted against true ratio in Figures 3-3 and 3-4.[4]

For true ratios of 5 to 1 or greater, percentage correct was considerably higher for words than for occupations;

[4] The tables on which these figures are based may be obtained from the authors.

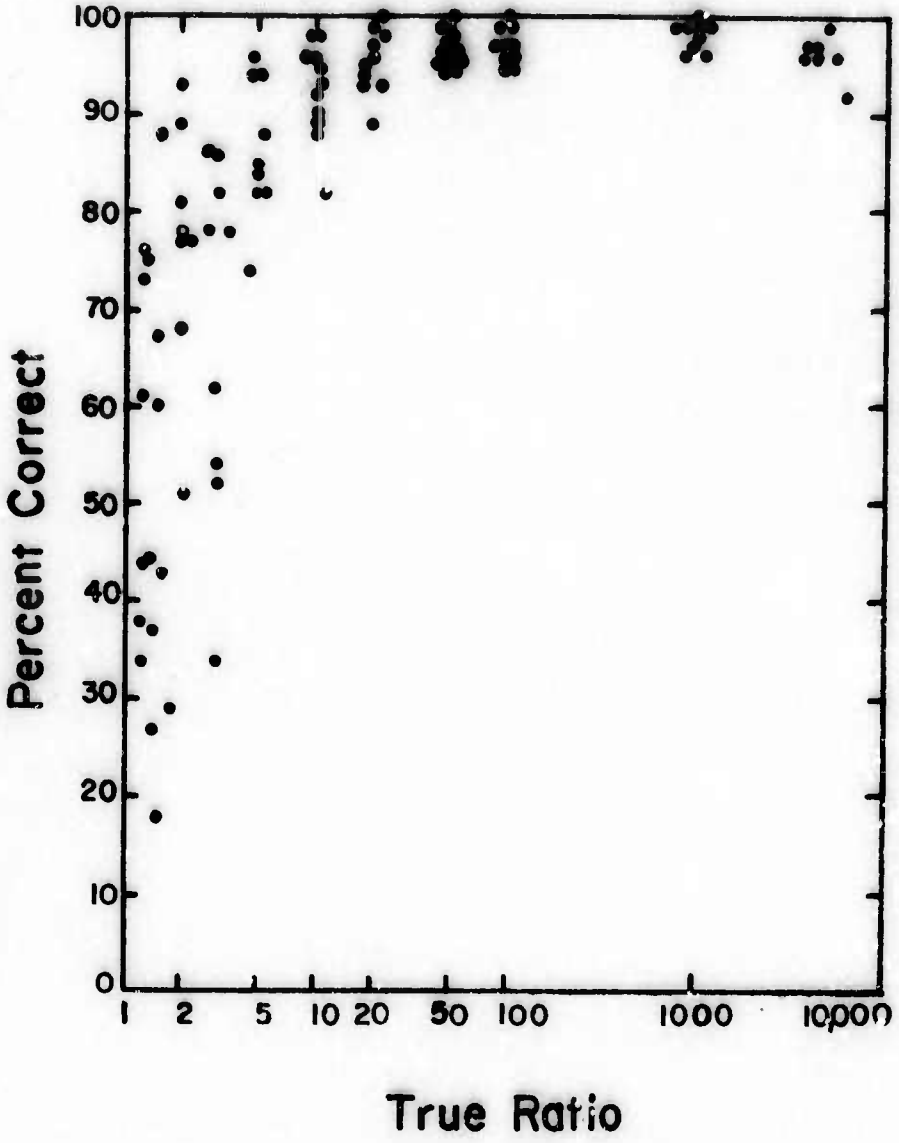


FIGURE 3-1. PERCENT CORRECT AS A FUNCTICN OF TRUE RATIO FOR 100 PAIRS OF WORDS

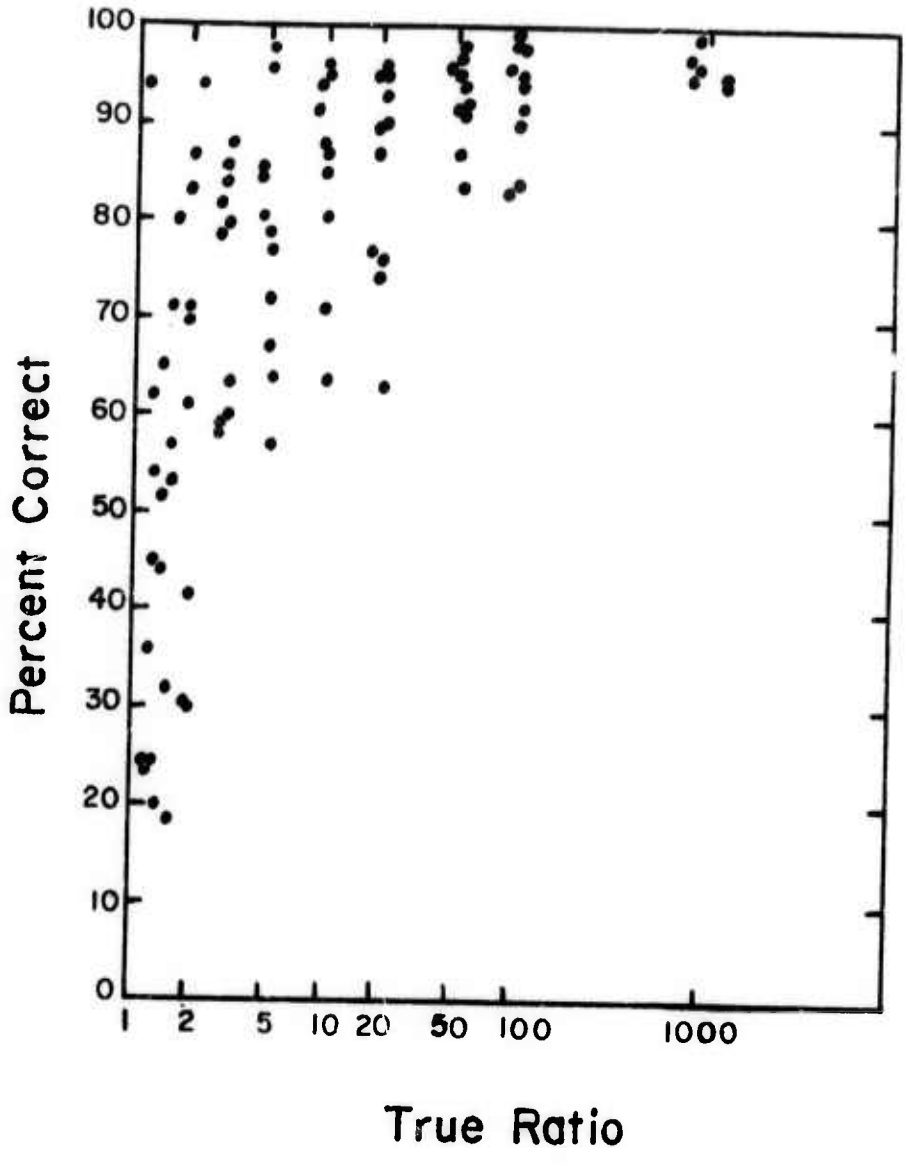


FIGURE 3-2. PERCENT CORRECT AS A FUNCTION OF TRUE RATIO FOR 95 PAIRS OF OCCUPATIONS

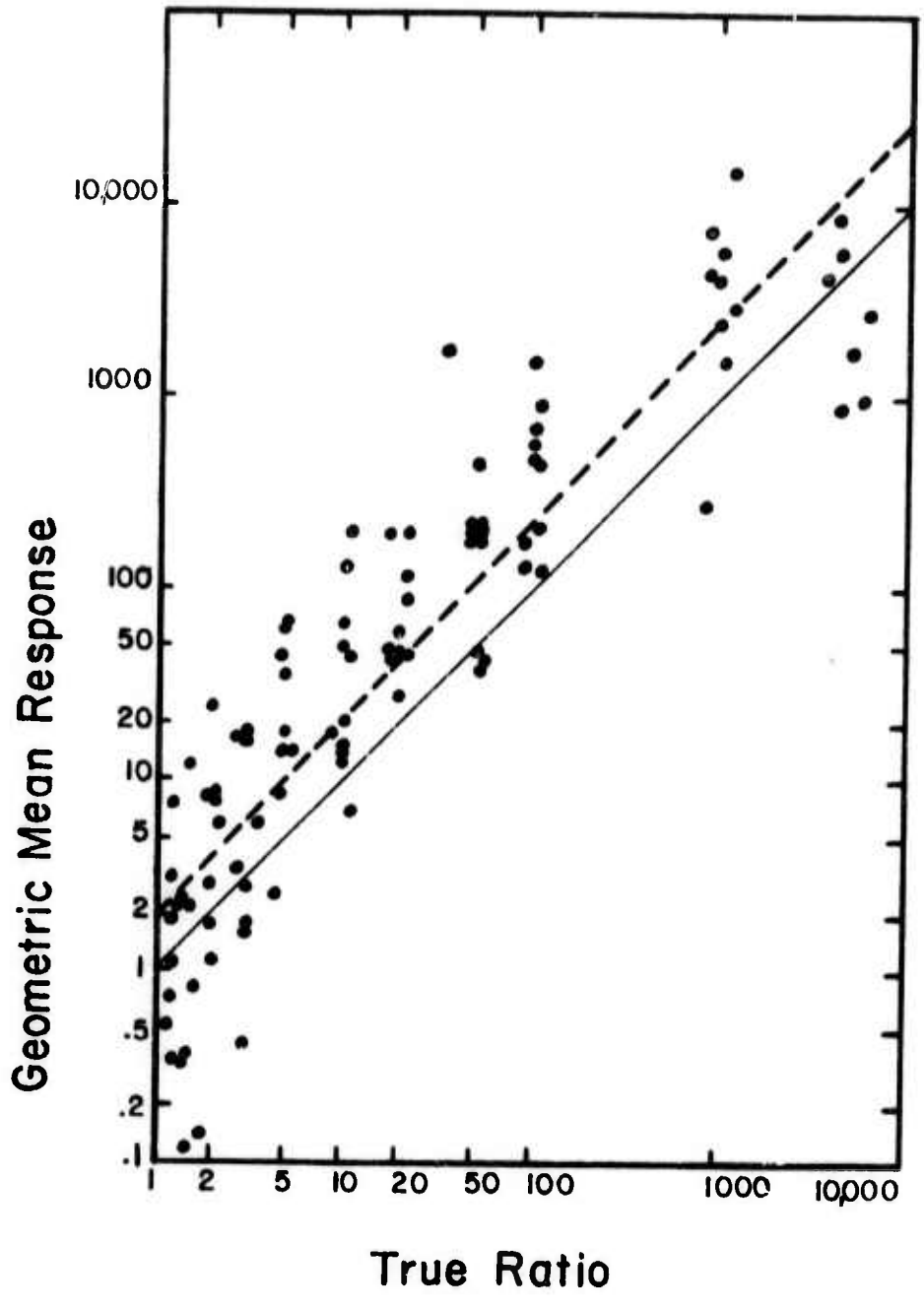


FIGURE 3-3. GEOMETRIC MEAN RATIO AS A FUNCTION OF TRUE RATIO FOR 100 PAIRS OF WORDS
DASHED LINE IS REGRESSION LINE

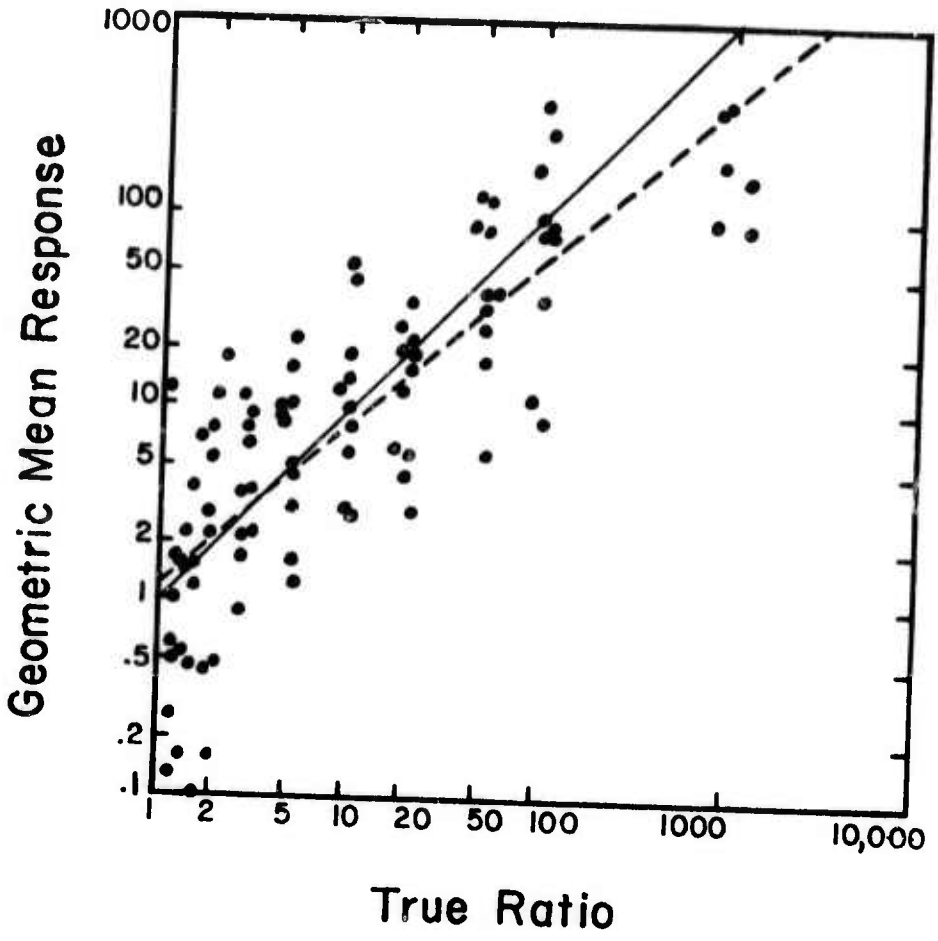


FIGURE 3-4. GEOMETRIC MEAN RATIO AS A FUNCTION OF TRUE RATIO FOR 95 PAIRS OF OCCUPATIONS
DASHED LINE IS REGRESSION LINE

below 5 : 1 there was no difference. For true ratios larger than 2 : 1, both words and occupations were more accurately discriminated than were causes of death (compare Figures 3-1 and 3-2 with Figure 2-1). For true ratios ≤ 2 : 1, there were again numerous errors of discrimination.

Geometric mean judged ratios for words and occupations were considerably closer to the corresponding true ratios than were judged ratios for causes of death, as may be seen by comparing Figures 3-3 and 3-4 (words and occupations) with Figure 2-3 (causes of death). The correlation between judged and true ratios was higher for words (.90) than for occupations (.81), but since the scatter about the regression line is not notably greater, this effect may be attributed to the greater range of true ratios for words.

The regression equations for the two causes-of-death groups and for words and occupations are shown in Table 3-3.[5] The slope for occupations was somewhat flat, but words showed a slope near unity which, taken with the intercept of 1.95, indicated a systematic tendency toward overestimation.[6]

---

[5] The regressions are linear in log-log space. However, the antilogs of the intercepts are shown here. These are the predicted judged ratios associated with a true ratio of 1.00.

[6] Carroll (1971), who elicited direct (magnitude) estimates of 60 words (12 of which were used here), found a correlation of .92 between assessed and actual values. His regression line had a slope of .58.

TABLE 3-3

REGRESSION EQUATIONS FOR GEOMETRIC MEAN JUDGED RATIO AGAINST TRUE RATIO

	Slope	Intercept	r
Causes of death: Students	.57	1.40	.69
Causes of death: LWV	.70	2.03	.75
Words	1.03	1.95	.90
Occupations	.84	1.17	.81

3.2.2 Consistency. The consistency of subjective ordering of the stimuli was sought by analyzing the triads in the words and occupations pairs. Of the 39 triads contained in the words task, 28 showed strong stochastic transitivity, 10 showed moderate stochastic transitivity, and one was intransitive. The one intransitive triad involved three pairs for which the subjects were quite indecisive (57% of subjects thought "in" was more likely than "that"; 56% "that" more likely than "for"; and 51% "for" more likely than "in"). Of the 20 triads contained in the occupations task, 17 showed strong stochastic transitivity and 3 showed moderate stochastic transitivity.

The log geometric mean ratio response to the third pair of each triad was correlated with the log of the product of the responses of the other two pairs; these correlations were .94 for words (slope 1.21, antilog of intercept = .80) and .76 for occupations (slope .64, antilog of intercept = 5.32). Thus, words and occupations judgments showed considerable internal consistency, as found with causes of death.

3.2.3 Comparison with Experiment 1. The purpose of Experiment 2 was to find out whether the major findings of Experiment 1 were specific to lethal events. Three results of this comparison are noteworthy. First, subjects responded more accurately to words than to occupations; causes of death were worse yet. This may be due to exposure: we experience many more samples of English text each day than examples of people working in occupations, and our exposure to death is even more limited. Another possible reason for poorer performance with causes of death is that our exposure to these events is systematically biased. We shall discuss this bias later in the paper.

Second, we found that causes-of-death subjects tended to underestimate large ratios. This tendency did not appear with words, but was found with occupations: the six occupation pairs with the highest ratios were all underestimated by at least a factor of two. Given these conflicting results, it is difficult to ascertain the generality of the tendency towards underestimation of high-ratio events.

Third, we found strong evidence in these new tasks that subjects possess consistent subjective frequency scales for these content areas, as they did for causes of death.

4.0 EXPERIMENT 3: DIRECT ESTIMATES OF EVENT FREQUENCIES

Experiment 1 suggested that subjects have a consistent underlying scale for the frequency of lethal events, although that scale deviates markedly from the statistically correct one. Unfortunately, the incomplete paired-comparison design used in Experiment 1 did not permit the subjective scale to be uncovered for all events. When the judged relative frequencies for a given pair were in error, it was difficult to determine whether judgments were biased for one, the other, or both members of the pair. Experiment 3 elicited direct estimates to clarify the nature of the biases for individual lethal events.

4.1 Method

The subjects were 74 respondents to an advertisement in the University of Oregon campus newspaper. Each subject was assigned to one of two groups. One group (N = 40) was told that the frequency of deaths in the U.S. due to Motor Vehicle Accidents was 50,000 per year (Group MVA). Using this value as a standard, they were asked to estimate the frequency for the other 40 lethal events shown in Table 2-1. The remaining 34 subjects (Group E) were given Electrocution = 1000 as a standard. The glossary used in Experiment 1, which defined some of the events, was provided. The 41 events were listed in alphabetical order on a single sheet. Subjects were encouraged to erase and change answers to make the relative frequencies of the entire set consistent with their best opinions.

Since there were about 205,000,000 persons in the United States when the data were collected, the rates per 10^8 shown in Table 2-1 were multiplied by 2.05 to provide statistical frequency against which to compare subjects' judgments. The standards given to the subjects, 1000 for electrocutions and 50,000 for motor vehicle accidents, were close to these computed statistical frequencies (1025 and 55,350, respectively).

4.2 Results

The data for one subject from Group MVA and two subjects from Group E were excluded from all analyses because they gave unreasonably high estimates (the sum of their estimates for all 41 causes of death exceeded 50,000,000, whereas the sum of the statistical frequencies is 3,553,004). Another subject was excluded from Group E because of unusually low responses. All of this subject's responses were below 1000 (the value of the standard); 38 of 40 responses were less than 100. As a result of these exclusions, the data presented below are based on 39 subjects in Group MVA and 31 subjects in Group E.

Because arithmetic means tend to be unduly influenced by occasional extreme values, the present results are based on the geometric means of the estimates. The use of medians leads to essentially the same results. For both groups, the correlation between log geometric mean and log median was $r = .99$ (for Group MVA, slope = 1.01, antilog of intercept = .97; for Group E, slope = 1.00, antilog of intercept = 1.17).

TABLE 4-1

RESULTS FROM DIRECT ESTIMATES

	Rate per 2.05×10^8	MVA		Electrocution	
		Geom. Mean	Ratio of Judged to Predicted	Geom. Mean	Ratio of Judged to Predicted
Smallpox	0	88		37	
Poison by Vitamin	1	237	1.27	44	1.16
Botulism	2	379	1.97	88	1.96
Measles	5	331	1.39	85	1.47
Fireworks	6	331	1.54	77	1.26
Smallpox Vaccination	8	38	.17	14	.22
Whooping Cough	15	171	.69	51	.62
Polio	17	202	.80	47	.55
Venomous Bite or Sting	48	535	1.67	233	1.85
Tornado	90	688	1.82	463	2.86
Lightning	107	128	.32	64	.37
Non-venomous Animal	129	298	.71	102	.54
Flood	205	863	1.77	627	2.71
Excess Cold	334	468	.81	211	.73
Syphilis	410	717	1.15	338	1.05
Pregnancy, etc.	451	1,932	2.98	935	2.78
Infectious Hepatitis	677	907	1.19	328	.80
Appendicitis	902	880	1.03	416	.87
Electrocution	1,025	586	.65	1,000*	1.96
Motor/Train Collision	1,517	793	.74	598	.95
Asthma	1,886	769	.65	333	.47
Firearms	2,255	1,623	1.26	1,114	1.42
Poisoning	2,563	1,318	.96	778	.92
Tuberculosis	3,690	966	.59	448	.43
Fire and Flames	7,380	3,814	1.62	2,918	1.86
Drowning	7,380	1,989	.85	1,425	.91
Leukemia	14,555	2,807	.81	2,220	.92
Accidental Falls	17,425	2,585	.68	2,768	1.03
Homicide	18,860	8,441	2.10	3,691	1.30
Emphysema	21,730	3,009	.69	2,696	.86
Suicide	24,600	6,674	1.42	3,280	.97
Breast Cancer	31,160	3,607	.66	2,436	.61
Diabetes	38,950	2,138	.34	1,019	.22
Motor Vehicle Accident	55,350	50,000*	6.34	33,884	5.76
Lung Cancer	75,850	9,723	1.00	9,806	1.33
Stomach Cancer	95,120	4,878	.43	2,209	.26
All Accidents	112,750	86,537	6.77	91,285	9.32
Stroke	209,100	10,668	.54	4,737	.31
All Cancer	328,000	47,523	1.70	43,772	2.00
Heart Disease	738,000	25,900	.49	21,503	.51
All Disease	1,740,450	80,779	.75	97,701	1.14

* Standard

The rank orders of the geometric means for the direct estimates were quite similar across the two subject groups (r = .98 for the log geometric means). However, as shown in Table 4-1, the geometric means for the MVA group were larger than those for Group E for 34 of 41 causes (sign test; $p < .001$). This difference may be due to MVA subjects anchoring on a larger standard than that presented to E subjects. (The two columns in Table 4-1 labeled Ratio of Judged to Predicted will be discussed later in the paper.)

4.2.1 Accuracy. Figures 4-1 and 4-2 show the geometric mean judgments plotted against the statistical rates (excluding smallpox). The best-fitting quadratic curves are also shown. For both groups, quadratic equations provided a significantly better fit ($p < .01$) to the data than linear equations. The equations for the quadratic curves, the correlations between the observed data and the results predicted from these curves, and the linear correlations are all given in Table 4-2.

TABLE 4-2

QUADRATIC FIT TO THE DIRECT ESTIMATES DATA

Quadratic Equation	R	r
Group MVA: $\log GM = .07 (\log TF)^2 + .03 \log TF + 2.27$	.92	.89
Group E: $\log GM = .05 (\log TF)^2 + .22 \log TF + 1.58$	.93	.91

GM: Geometric mean response
TF: True frequency
R: Quadratic correlation
r: Linear correlation

For both groups, low frequency events were overestimated, while high frequency events were underestimated. As shown by the quadratic curve in Figure 4-1, the crossover point for Group MVA was at a true rate of about 800; all events with frequencies lower than that were overestimated, while all above that point were underestimated. For Group E (see Figure 4-2) the crossover point was less clear; it occurred around a true rate of 250.

4.2.2 Secondary bias. Deviations from the regression curves were quite similar in Figures 4-1 and 4-2. The correlation between the two groups' residual values (i.e., the vertical distance between each point and the regression curve) was .91 across the 40 items (excluding smallpox), indicating a consistent secondary bias above and beyond the primary bias (overestimation of low frequencies and underestimation of high frequencies) evidenced by the regression curves. The antilogs of these residuals are shown in Table 4-1, in the columns labeled "Ratio of Judged to Predicted." Some of the items with large residuals are labeled on the two figures. The similarity between the two groups of subjects, relative to their own regression lines, is striking. Frequency of death due to all accidents, motor vehicle accidents, pregnancy, flood, tornado and cancer was relatively overestimated by both groups. Death due to smallpox vaccination, diabetes, lightning, heart disease, tuberculosis and asthma was relatively underestimated.

4.2.3 Comparison with Experiment 1. Overall, there is a close relationship between the direct estimates of the present experiment and the paired-comparison results of Experiment 1. From the geometric means of the direct

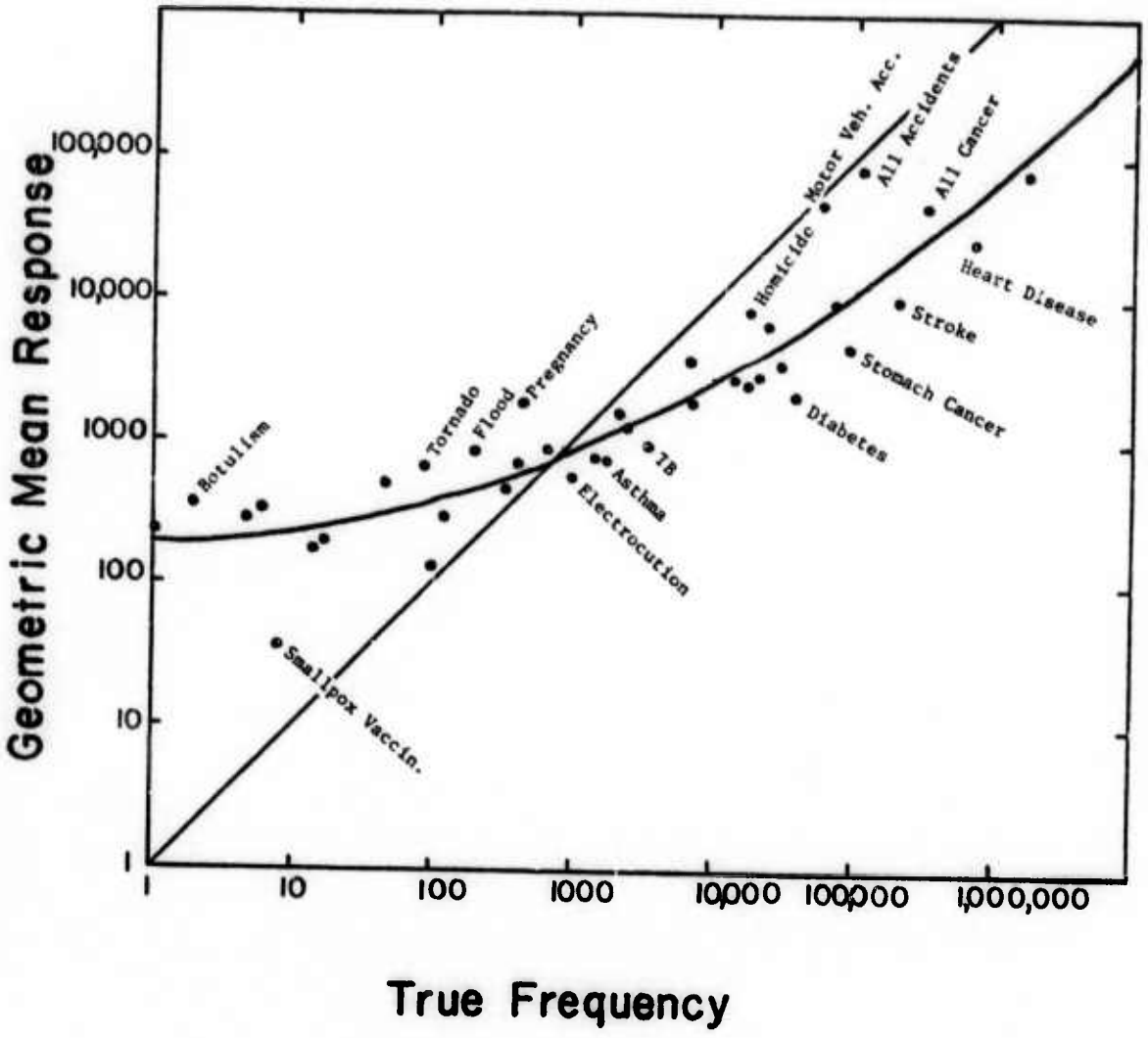


FIGURE 4-1. GEOMETRIC MEAN DIRECT ESTIMATES OF FREQUENCY AS A FUNCTION OF TRUE FREQUENCY: MOTOR VEHICLE ACCIDENT (MVA) GROUP

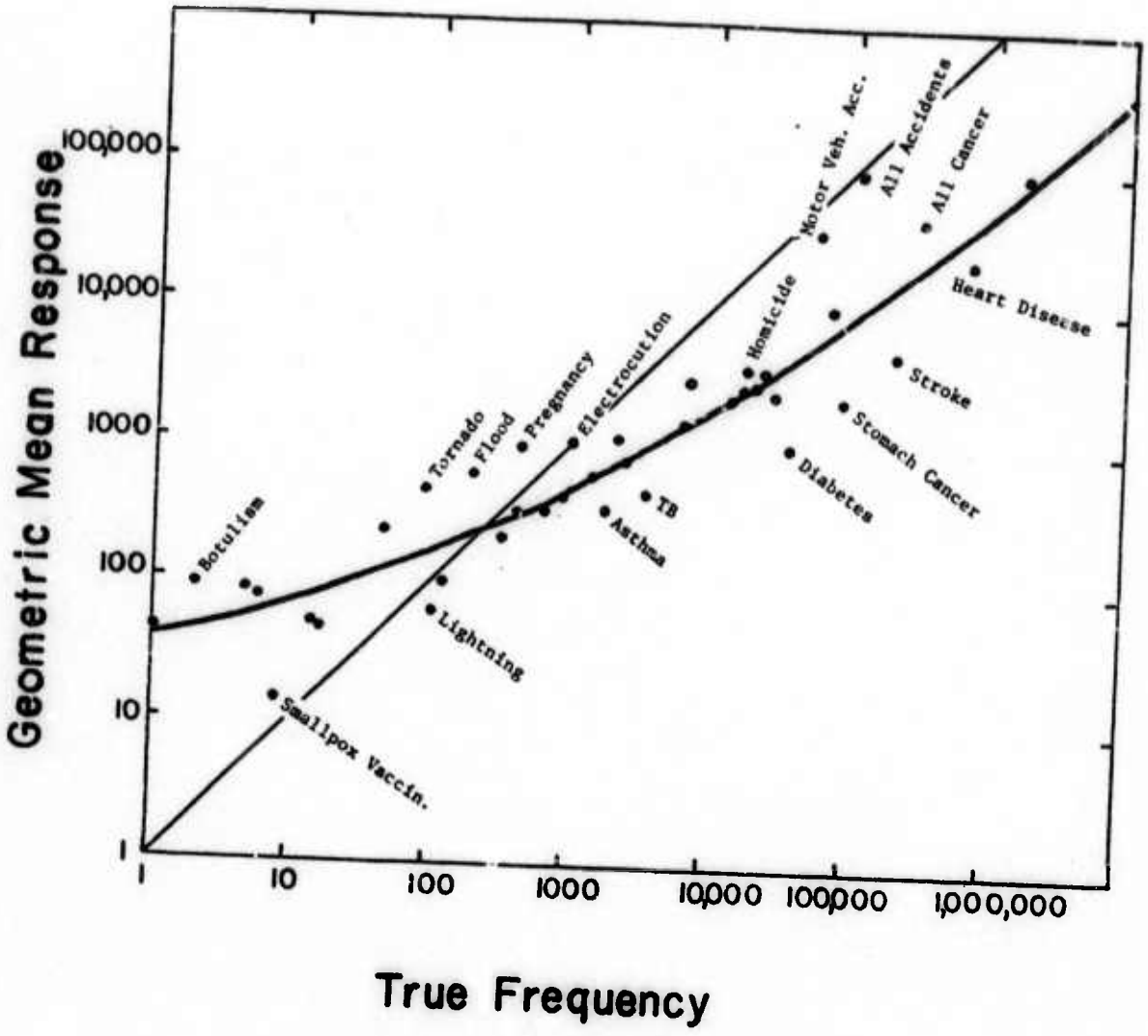


FIGURE 4-2. GEOMETRIC MEAN DIRECT ESTIMATES OF FREQUENCY AS A FUNCTION OF TRUE FREQUENCY: ELECTROCUTION (E) GROUP

estimates one can compute ratios for each of the 106 pairs studied in Experiment 1. The logs of these derived ratios were highly similar to the logs of the geometric mean frequency ratios from Experiment 1 (college students): $\underline{r}$ = .94 for the MVA group and .93 for the E group (across all 106 pairs).

Neither the judged ratios from Experiment 1 nor the ratios derived from the direct estimates of the present experiment were consistently closer to the true ratios. The judged ratios from Experiment 1 were less accurate when the true ratio was low (< 10 : 1) and more accurate when the true ratio was high (≥ 10 : 1).

4.2.4 Individual differences. For each subject the linear correlation between log response and log true rate was calculated across the 40 stimuli (excluding smallpox). Linear correlations were used after visual examination of the data plots revealed that only a few of the subjects showed the curvilinearity found in the group results. Group E showed a range from .61 to .92 and a median of .77. Within Group MVA, correlations ranged from .28 to .90; median .66.

5.0 EXPERIMENT 4: EXPERIENCE AND BIAS

Experiments 1 and 3 demonstrated that the frequencies of some lethal events are consistently misjudged. In hopes of learning more about the nature of these errors and biases, Experiment 4 examined people's direct and indirect experiences with these events and some of the events' special characteristics. Eight different characteristics were assessed for each lethal event and then used to predict the errors found in Experiments 1 and 3. Four of the measures assessed how much experience subjects have had with the different causes of death. Two measures reflected the frequency with which causes of death appear in newspaper articles. The final measure reflected the degree to which the various causes of death were perceived as being catastrophic (inflicting simultaneous multiple casualties) and lethal (inevitably producing death for people suffering from the condition).

5.1 Method

5.1.1 Experience ratings. A new group of 61 subjects recruited through the campus newspaper was asked to rate each of the 41 causes of death according to their personal experiences with the event as a cause of death and suffering.

Two ratings of indirect experience were obtained by asking subjects to indicate how often they had heard about the event via the news media (newspapers, magazines, radio, television, etc.) as (a) a cause of death and (b) a cause of suffering (but not death). Ratings were made on a five-point

scale whose extreme categories were "never" (coded as 1) and "often" (coded as 5).

Subjects' direct experience with the 41 events as causes of death were elicited by having them check one of the following three statements for each event:

> Code 3: At least one close friend or relative has died from this.
>
> Code 2: Someone I know (other than a close friend or relative) has died from this.
>
> Code 1: No one I know has died from this.

Direct experience with these events as causes of suffering was elicited with similar questions, with the word "died" replaced by the phrase "suffered (but not died)".

Thus, each subject provided four ratings for each of the 41 events. These were ratings of:

(a) indirect death (coded 1 to 5),
(b) indirect suffering (coded 1 to 5),
(c) direct death (coded 1 to 3), and
(d) direct suffering (coded 1 to 3).

5.1.2 Newspaper coverage. The news media provide two kinds of information about causes of death. One, as noted earlier, is reports of the latest statistical analyses (Figure 1-1). The other, far more prevalent, is the day-to-day reporting of fatalities, as they happen. The latter is likely to be biased towards violent and catastrophic events (see, for

example, Arlen's [1975] survey of television's treatment of death). Because of the potential importance of media exposure, we supplemented people's ratings of their indirect (media) experiences with a survey of newspaper reports. The local daily newspaper (the Eugene Register Guard) was examined on all days of alternative months for a year, starting with January 1, 1975 (for a total of 184 days). Two tallies were made for each cause of death: the total number of deaths reported and the total square inches of reporting devoted to the deaths (excluding photographs).

5.1.3 Catastrophe ratings. Economist Theodore Bergstrom (1974) has asked whether catastrophic events, with multiple victims in close geographic and temporal proximity, will be judged as more likely than events which take as many lives but in a less spectacular, one-at-a-time fashion. He hypothesized that catastrophes are more spectacular and thus more memorable, a speculation in keeping with availability considerations. On the other hand, the more frequent instances of non-catastrophic events may lead them to be perceived more accurately, while casualties from catastrophic events may be underestimated because of their massed presentation (Hintzman, 1976).

To assess catastrophic potential, 13 employees of the Oregon Research Institute were asked to estimate the average number of people who die from a single fatal episode of each of the 41 causes of death.

5.1.4 Conditional death ratings. In Experiments 1 and 3, subjects appeared to underestimate (relative to the

regression line) the frequencies of deaths due to events that are common in non-fatal form, such as smallpox vaccination and asthma. One possible explanation of this error is that subjects both confused P(A|B) with P(B|A) and failed to appreciate the importance of base rates (Tversky & Kahneman, 1974; Bar-Hillel, 1977). Consider the question of whether a randomly selected death is most likely to be due to smallpox or smallpox vaccination. This question calls for comparing P(smallpox|death) with P(smallpox vaccination|death), the latter being statistically greater. However, subjects may be relying on P(death|smallpox) and P(death|smallpox vaccination) to answer such questions. If the base rates for the various events are discrepant (as they are in this case), the resulting judgments will be in error.

To explore the role of this characteristic, 31 college students were asked to rate the probability of death given that one suffered from or experienced each condition. The ratings were made on a scale from 0 ("surely won't die") to 20 ("surely will die").

5.2 Results

5.2.1 Mean values. Mean values for the six subjective scales and the two newspaper measures are shown in Table 5-1.

As one would expect, subjects reported greater experience with these events as causes of suffering than as

TABLE 5-1

RATINGS ON EIGHT PREDICTOR VARIABLES

	Indir. Death	Indir. Suffer	Direct Death	Direct Suffer	News-paper Freq.	News-paper Inches	Catas-trophe	Condition-ality
Smallpox	2.20	2.48	1.02	1.33	0	0	1.35	5.87
Poisoning by vitamins	1.23	1.43	1.00	1.07	0	0	1	4.36
Botulism	2.82	2.82	1.03	1.36	0	0	1.49	10.32
Measles	2.07	3.00	1.05	2.41	0	0	1	1.81
Fireworks	2.43	2.85	1.10	1.56	0	0	1	3.73
Smallpox vaccination	1.30	1.71	1.03	1.41	0	0	1	.71
Whooping cough	1.48	1.95	1.00	1.38	0	0	1	3.84
Polio	2.49	2.87	1.15	1.77	0	0	1	4.81
Venomous bite or sting	2.41	2.97	1.05	2.15	0	0	1	6.84
Tornado	3.46	3.75	1.07	1.38	36	153.5	4.51	6.25
Lightning	2.34	2.38	1.05	1.23	1	.8	1.01	10.06
Non-venomous animal	2.30	2.89	1.03	1.82	4	33.8	1	3.19
Flood	3.66	4.05	1.12	1.56	4	41.8	5.57	6.52
Excess Cold	2.62	2.93	1.15	1.57	0	0	1.20	10.15
Syphilis	2.51	3.67	1.07	1.79	0	0	1	5.19
Pregnancy, etc.	3.07	3.84	1.13	2.03	0	0	1.01	4.57
Infectious hepatitis	2.03	2.77	1.12	2.02	0	0	1	7.79
Appendicitis	2.00	2.67	1.10	2.30	0	0	1	3.53
Electrocution	2.90	2.69	1.21	1.57	5	42.2	1	15.81
Motor/train collision	3.03	2.85	1.23	1.28	0	0	2.12	14.87
Asthma	1.62	3.13	1.18	2.41	1	1.9	1	2.07
Firearm accident	3.89	3.87	1.44	1.67	8	28.2	1.02	10.34
Poisoning solid/liquid	3.02	3.05	1.10	1.61	3	17.9	1.03	10.81
Tuberculosis	2.71	3.13	1.10	1.61	0	0	1.08	7.68
Fire and flames	4.07	4.15	1.20	1.71	94	320.7	1.73	10.58

TABLE 5-1 (CONT'D.)

	Indir. Death	Indir. Suffer	Direct Death	Direct Suffer	News-paper Freq.	News-paper Inches	Catas-trophe	Condition-ality
Drowning	3.82	3.23	1.69	1.68	47	247	1.07	17.65
Leukemia	3.56	3.38	1.36	1.23	1	14.8	1	15.00
Accidental falls	3.18	3.54	1.31	2.43	15	124.8	1.03	4.79
Homicide	4.69	4.33	1.39	1.23	278	5042.9	1.06	18.32
Emphysema	3.02	3.36	1.31	1.75	1	1.1	1	11.03
Suicide	4.00	3.66	1.74	1.71	29	356.7	1	17.23
Breast cancer	3.03	4.33	1.38	2.00	0	0	1	9.39
Diabetes	2.37	3.49	1.31	2.39	0	0	1	6.45
Motor vehicle accident	4.69	4.71	2.03	2.61	298	1440.5	1.64	8.97
Lung cancer	4.15	4.21	1.82	1.66	3	35.9	1	14.26
Stomach cancer	2.89	3.08	1.59	1.59	0	0	1	11.87
All accident	4.44	4.64	2.05	2.43	715	2861.4	1.70	6.97
Stroke	3.87	3.98	1.95	2.18	12	130.7	1	11.76
All cancer	4.54	4.59	2.38	2.34	25	188.5	1	13.16
Heart disease	4.28	4.34	2.15	2.10	49	303.4	1	13.00
All disease	4.48	4.49	2.25	2.44	111	727.1	1.19	8.00
Range of scale	1-5	1-5	1-3	1-3	1-∞	1-∞	1-∞	0-20
Mean	3.04	3.35	1.35	1.80	42.4	295.5	1.31	8.77

causes of death. The most frequently experienced event was motor vehicle accidents, while the lowest ratings were given to poisoning by vitamins.

During 184 days of newspaper reporting, 19 of the listed causes of death were never mentioned. Some of these 19 causes are quite frequent: cancer of the digestive system, diabetes, breast cancer and tuberculosis. In contrast, the eighth most frequently reported cause of death in the newspapers, tornadoes, is in fact relatively rare. The reported tornado deaths may represent all deaths from this cause in the United States during the dates covered. Note also that homicide, which is 23% less frequent than suicide, was reported 9.6 times as often, with 15 times as much space devoted to it.[7]

Few of the listed causes of death can be classed as catastrophic in terms of the perceived number of people dying on a single occasion. Flood, tornado and motor vehicle/train collisions led the catastrophe ratings.

The conditional death ratings seem reasonable. The lowest rating was given to smallpox vaccination, while the highest was to homicide, followed by drowning. Some chronic diseases, asthma, diabetes, syphilis and tuberculosis, were rated below the overall mean of 8.77, but emphysema (11.03) and heart disease (13.00) were both rated well above the mean.

---

[7] This result may be even more extreme than it appears, since there is good reason to suppose that the official records we used to establish "true" rates underestimate the frequency of suicide.

5.2.2 Correlations: paired comparisons. Correlational analyses were performed to determine whether the eight measures predict the judgments and biases found in Experiments 1 and 3. In order to predict the paired-comparison results, a difference score was formed on each measure for each of the 101 pairs (excluding smallpox) by subtracting the score associated with the less likely cause of death from the score associated with the more likely cause of death.

Two aspects of paired-comparison data were predicted from these difference scores: (a) the log geometric mean response to the 101 paired items (excluding smallpox), and (b) the index of secondary bias used in Experiment 1 (the signed difference between the log geometric mean of the judged likelihood ratios and the log geometric mean predicted by the regression lines shown in Figures 2-3 and 2-4).

Table 5-2 shows the intercorrelation matrix for the four response variables (log geometric mean ratio judgments and residuals for students and for League members), the true ratio, and the eight predictor variables.

The lower left rectangle of correlations indicates the predictive power of the eight independent variables. Three of the four experience ratings showed strong correlations with the four response variables. Note that these ratings correlated more highly with the subjects' responses than with the true ratios. Only the ratings of direct suffering showed low correlations with subjects' responses.

News frequency and news inches were also modestly good predictors of the response variables, even though they were

TABLE 5-2

PAIRED COMPARISON CORRELATION MATRIX

	Students Log Geom. Mean	LWV Log Geom. Mean	Student Residuals	LWV Residuals	Log True Ratio	Indirect Death	Indirect Suffering	Direct Death	Direct Suffering	News Frequency	News Inches	Catastrophe	Conditional Death
Students LGM	1.00												
LWV LGM	.95	1.00											
Student Resid.	.73	.59	1.00										
LWV Resid.	.66	.66	.90	1.00									
Log True Ratio	.69	.75	.00	.00	1.00								
Ind. Death	.76	.67	.66	.55	.41	1.00							
Ind. Suffer	.74	.72	.57	.55	.48	.78	1.00						
Dir. Death	.83	.78	.55	.47	.62	.59	.48	1.00					
Dir. Suffer	.17	.26	.01	.08	.26	-.19	.27	.23	1.00				
News Freq.	.44	.34	.58	.4	.03	.38	.39	.28	.15	1.00			
News Inches	.41	.34	.49	.44	.07	.42	.34	.19	-.09	.77	1.00		
Catastrophe	.11	.04	.25	.17	-.10	.38	.40	-.07	-.16	.13	-.01	1.00	
Cond. Death	.37	.31	.19	.09	.34	.61	.16	.35	-.56	-.05	.22	-.02	1.00

not correlated with true ratio. This lack of correlation with true ratio demonstrates the biased view of reality that newspapers present.[8] The catastrophe ratings showed quite low correlations with all other variables. This may be due, in part, to the lack of variance in these ratings; over half were equal to 1.0, and only 10 of 41 were greater than 1.08. Thus, most of the 101 differences formed from these ratings were at or near zero. Finally, conditional death ratings were slightly correlated with the geometric mean responses, but not with the residuals.

The correlations among the eight predictor measures are also shown in Table 5-2. Indirect death, indirect suffering, and direct death ratings showed fairly high intercorrelations, but low correlations with direct suffering. The two newspaper measures were highly intercorrelated. However, these newspaper measures correlated only moderately (.38, .42) with the indirect death ratings, even though the instructions for the latter task emphasized newspaper coverage.

Subjects' paired-comparison judgments correlated with the frequencies of newspaper coverage, which we know to be biased. Therefore, we might expect that ratings of direct experience (which might be less biased) would provide more accurate estimates of the true ratios than did the judgments of frequency. However, this does not turn out to be a successful debiasing technique. Although the direct death rating correlated more highly with the true ratio (r = .62)

---

[8] Comparable evidence of bias in another newspaper may be found in Combs and Slovic, Note 2.

than did any of the other predictor measures, the mean paired comparison judgments did even better (r = .68 and .75), despite being contaminated by various biases. Thus our subjects' frequency judgments contained valid information transcending their aggregate direct experiences.

5.2.3 Correlations: direct estimates. Parallel analyses were performed for the direct estimates of causes of death collected in Experiment 3.

The correlation matrix for these data is shown in Table 5-3. The first two variables are the log geometric means for the two groups of subjects, those given Motor Vehicle Accidents as a standard (Group MVA) and those given Electrocution as a standard (Group E). The next two variables are the residuals computed from the quadratic curves fit to the two groups' data (these residuals are the logs of the measures called "Ratio of Judged to Predicted" in Table 4-1). Following these four variables are the log true frequency for the causes of death and the eight predictor measures. All correlations were computed across the 40 lethal events excluding smallpox.

All four experience ratings (direct and indirect suffering and death) were highly correlated with the subjects' geometric mean responses. The correlations between the experience ratings and the true frequency were somewhat lower. The ratings were only moderately correlated with the residuals of the subjects' responses from the regression line. The two newspaper measures showed predictive power for both the responses and the residuals. Catastrophe ratings showed

TABLE 5-3

DIRECT ESTIMATES CORRELATION MATRIX

	MVA Log Geom. Group Mean	E Log Geom. Group Mean	MVA Group Residuals	E Group Residuals	Log True Frequency	Indirect Death	Indirect Suffering	Direct Death	Direct Suffering	News Frequency	News Inches	Catastrophe	Conditional Death
MVA LGM	1.00												
E LGM	.98	1.00											
MVA Resid.	.40	.35	1.00										
E Resid.	.36	.38	.91	1.00									
Log True Freq.	.89	.91	.00	.00	1.00								
Ind. Death	.85	.86	.45	.48	.74	1.00							
Ind. Suffer	.86	.86	.46	.44	.76	.88	1.00						
Dir. Death	.90	.88	.19	.19	.82	.77	.71	1.00					
Dir. Suffer	.52	.50	.22	.16	.46	.21	.49	.47	1.00				
News Freq.	.56	.54	.59	.56	.33	.50	.47	.45	.29	1.00			
News Inches	.45	.41	.45	.36	.29	.50	.43	.31	.04	.77	1.00		
Catastrophe	-.03	.02	.29	.40	-.12	.21	.21	-.11	-.17	.09	.02	1.00	
Cond. Death	.47	.51	.04	.08	.54	.65	.37	.47	-.28	.10	.30	-.07	1.00

weak correlations with the residuals and none with the geometric mean responses, while conditional death ratings correlated with the geometric mean responses, but not with the residuals.

As with the paired-comparison data (Table 5-2), the direct death rating correlated most highly of the eight measures with the true frequency (r = .82). However, it could not successfully be substituted for the direct estimates of frequency in an attempt to improve accuracy, since these direct estimates correlated .89 and .91 with the "true" ratios. Again, subjects' frequency judgments reflected something valid beyond their direct experiences.

The intercorrelations among the predictor variables shown in the right triangle in Table 5-3 are necessarily similar to those shown in Table 5-2, since they are based on the same data (expressed there as differences between pairs).

5.2.4 Regression analyses predicting responses and biases. To bring greater clarity to this mass of correlations, eight stepwise regressions were performed. Four of these analyses predicted the log geometric mean responses of the four separate groups of subjects: students' paired-comparisons, League members' paired-comparisons, Group E's direct estimates, and Group MVA's direct estimates. The other four stepwise regression analyses predicted secondary bias (the residuals from the correlations of each of these four groups with the statistical frequencies).

The predictor variables for each of the stepwise regressions were the eight measures previously described,

using differences between 101 pairs to predict the paired-comparison data, or 40 mean ratings to predict the direct estimates and their residuals.

Because of the instability of stepwise regression solutions with highly intercorrelated predictors, our primary criterion for variable selection was replicability. Only variables that entered the equations for both League and student subjects in Experiment 1 or both Group E and Group MVA in Experiment 2 are discussed. Table 5-4 lists the variables that emerged from both groups of subjects. The inclusion criterion was an F to enter[9] of 3.0 or greater. The log geometric means were highly predictable, with multiple R's ranging from .88 to .96 using just three of the eight predictors. The residuals were also predictable, with multiple R's ranging from .64 to .80 using the variables selected by the stepwise regression.

TABLE 5-4

VARIABLES EMERGING FROM STEPWISE MULTIPLE REGRESSIONS IN BOTH REPLICATIONS

Dependent Variables			
Log Geometric Mean		Residuals	
Paired Comparisons	Direct Estimates	Paired Comparisons	Direct Estimates
Indirect Suffering	Indirect Suffering	Indirect Death	News Frequency
Direct Death	Direct Death	Direct Death	Catastrophe
	News Frequency	Conditional Death[a]	

[a] Negative weight

[9] An "F to enter" tests the significance of the increase in the proportion of explained variance achieved by including an additional variable in the regression equation.

Two variables, indirect suffering and direct death, did most of the job of predicting the subjects' log geometric mean responses for both paired comparisons and direct estimates. The regressions on the residuals show a more mixed pattern. For the residuals from the paired comparisons data, three predictors were common to both the student and League data: indirect death, direct death and conditional death, the latter with a negative weight, due to its low correlation with the dependent variable and its high correlation with indirect death. For the prediction of residuals from the direct estimates, news frequency and catastrophe ratings were the only predictors that were significant in both groups. In view of the highly skewed distributions of these two measures, it is somewhat surprising to see them emerge as valid predictors. However, news frequency correlated with direct estimate residuals higher than any other single predictor. And of the seven catastrophe ratings of 1.5 or greater, six (all accidents, motor vehicle accidents, flood, botulism, tornado and fire and flames) were among the ten causes of death with the highest residuals (i.e., the ten most overestimated causes of death, relative to the regression line).

The above analyses indicate that measures tapping the availability of information about causes of death do a good job of predicting subjects' perceptions of the relative frequencies of these causes of death. Further, we have shown that the consistent errors people make (the secondary bias) can be predicted from subjects' experience with these events and from salient features such as their catastrophic nature.

6.0 EXPERIMENT 5: DEBIASING

Despite the fact that subjects' responses in Experiments 1 and 3 were often biased, Tables 5-2 and 5-3 revealed no better single predictor of statistical frequencies. The systematic nature of these biases suggests that they could be corrected statistically, by using the best-fit curves to remove the primary bias and by using knowledge of personal experience or media exposure to reduce the secondary bias. The primary bias seems quite easy to correct; the regression equation derived from one set of causes of death could reasonably be used to correct a similar, untested set. However, statistical correction of the secondary bias would be more difficult; each cause of death would require its own correction factor. A simpler, more direct approach would be to train subjects to avoid these errors. Experiment 5 was designed to explore the possibility of eliminating the secondary bias. Subjects were briefed on the prevalence and nature of the bias in order to determine whether this knowledge could help them to be more accurate judges of relative frequency.

6.1 Study 5A

6.1.1 Method. In Study 5A, subjects made paired comparisons for 31 of the 106 pairs of Experiment 1. Twenty-one of these pairs were severely misjudged in Experiment 1 (either the percentage correct was less than 60 or the geometric mean was off by a factor of 9 or more). The geometric means of the remaining 10 were estimated moderately well (within a factor of 1.5). The present study was

conducted with a college student population similar to that in Experiment 1 and with the same instructions except that one group, the "debiasing" group (N = 30), was given the following special information:

> Note: In a previous study of this kind we found that, for some pairs, the relative likelihoods were greatly misperceived. Sometimes the ratio of the more likely to the less likely item was judged to be much greater than it really was. In other cases the ratio was judged much too small or even in the wrong direction; that is, the less likely item was judged to be more likely.
>
> We believe that when people estimate these likelihoods, they do so on the basis of a) how easy it is to imagine someone dying from such a cause, b) how many instances of such an event they can remember happening to someone they know, c) publicity about such events in the news media, or d) special features of the event that make it stand out in one's mind.
>
> Reliance on imaginability, memorability, and media publicity, although often useful, can lead to large errors in judgment. When events are disproportionately imaginable or memorable, they are likely to be overestimated. When they are rather unmemorable or unpublicized or otherwise undistinguished, they are likely to be underestimated. Events such as ulcers that are common, but usually non-fatal, may also be underestimated because people tend to imagine or remember them in their non-fatal form.
>
> Try not to let your own judgments be biased by factors such as imaginability, memorability, or media publicity.

A control group (N = 22) also judged the 31 pairs without receiving any special instructions.

6.1.2 Results. Examination of percentage correct revealed no evidence for debiasing. The original subjects

were best on 9 pairs, the control subjects best on 12 pairs, and the debiasing group subjects were best on 10 pairs.

A further search for improvement in the data of Study 5A can be made by comparing the ratio judgments of these two new groups of subjects either with the true ratios (under the assumption that the instructions exhorted the subjects to come closer to the truth) or with the ratios predicted from the regression analysis of the original subjects (under the assumption that the instructions emphasized the nature of the secondary bias, not the primary bias). Under either comparison, no evidence for effective debiasing can be seen. For geometric means, when the comparison is made to true ratio, the original group was best on 12 pairs, the controls on 6 pairs, and the debiasing group on 13 pairs. When compared with the predicted ratios, the original group was best on 12 pairs, the control group on 7 and the debiasing group on 12. Looking only at the 21 pairs that were originally judged poorly, there is still no evidence of improvement in the debiased group. Even those pairs on which the debiasing group did best showed only modest improvement. For example, death by diabetes is 95 times more likely than death by syphilis. The debiasing group was "superior" in giving a geometric mean response of 9.7 rather than the original group's geometric mean of 2.4. Death by stroke is 102,000 times more likely than death by botulism. The value predicted by the regression analysis of the original subjects was 1002. Those original subjects showed a strong secondary bias; their geometric mean response was 106. The debiasing experimental group gave a mean response of 135.

6.2 Study 5B

6.2.1 Method. A second debiasing study was undertaken to provide subjects even more opportunity for using knowledge of the secondary biases to improve their performance.

The subjects, drawn from the same student population, were shown 19 pairs of events. The instructions indicated that each of these pairs had been seriously misjudged in an earlier experiment (which was the case). For each pair, the subjects were given the response from Experiment 1 and were asked to improve it, that is, to give a new response that they thought would be closer to the true ratio.

The instructions for a debiasing group of 29 subjects included a discussion of the presumed sources of error, illustrated with several examples showing the possible effects of personal experience, media publicity, imaginability, etc., on previous subjects' judgments. A control group of 27 subjects did not receive this additional discussion.

The instructions read as follows. Brackets indicate material shown only to the debiasing group.

> We recently studied the ability of University of Oregon students to judge the likelihood of various causes of death in the United States.
>
> For example, subjects were given a pair of events such as:
>
> A. Measles
> B. Tornado

They were asked: Which causes more deaths annually in the U.S., A or B? They were also asked to estimate how many times more likely the more frequent cause of death was compared to the less frequent of the two.

We found that, for some pairs, the relative likelihoods were greatly misjudged. Sometimes the ratio of the more likely to the less likely item was judged much too small or even in the wrong direction; that is, the less likely item was judged to be more likely.

[We believe that when people estimate these frequencies, they do so on the basis of a) how easy it is to imagine someone dying from such a cause, b) how many instances of such an event they can remember happening to someone they know, c) publicity about such events in the news media, or d) special features of the event that make it stand out in one's mind.]

[When events are disproportionately imaginable or memorable, they are less likely to be overestimated. When they are rather unmemorable or unpublicized or otherwise undistinguished, they are likely to be underestimated. Events such as accidental falls, that are common but usually non-fatal, may also be underestimated because people tend to imagine or remember them in their non-fatal form.]

On the following pages there are 19 pairings of death-producing events. The relative likelihood of the more common to the less common event was greatly misperceived in each of these pairs.

[We want to see whether you can reduce the magnitude of the errors for these pairs. To do this think about how factors such as media coverage or ease of imagining or remembering the event as a cause of death are likely to work to bias the judgments for each of the pairs.]

Here are some examples to illustate the task:

		Previous Answer	Your Answer
A.	Hepatitis	B 4.55	
B.	Drowning		

The average subject chose B as more likely and judged it to be 4.55 times more likely than A. Which would you choose and what ratio would you give?

Actually, the correct answer is B and the true ratio is 10.9 to 1. We see that the average subject overestimated Hepatitis relative to Drowning. [Maybe this is because of the special attention given by the media to Hepatitis, especially in relation to abuse of hypodermic needles.]

Try this one:

		Previous Answer	Your Answer
A.	Leukemia	A 1.30	
B.	Accidental Falls		

The average subject thought death from leukemia was 30% more common (ratio 1.30 to 1) than death from falls. However, death from falls is really 20% more frequent. So the correct answer is B with a ratio of 1.20. [The error may stem from the dramatic nature of leukemia and the greater amount of media publicity it receives, or it may stem from the fact that accidental falls are common but usually non-fatal.]

For a final example, consider:

		Previous Answer	Your Answer
A.	Poisoning by solid or liquid	A 5.26	
B.	Tuberculosis		

The average subject thought death by poisoning was 5.26 times more likely than death from tuberculosis. However, death from tuberculosis is really 44% more frequent than death from poisoning so the correct answer is B with a ratio of 1.44. [Again, it is easy to see how media publicity regarding poisoning and the dramatic nature of the event could cause subjects to overestimate it compared to the drab, undramatic, perhaps old-fashioned disease, tuberculosis.]

Note that a ratio of 1.20 means 20% more likely,
1.50 means 50% more likely,
1.80 means 80% more likely, etc.

For each pair, write the letter of the item you think is a more likely cause of death and give your judgment about how many times more frequent the more frequent item is.

6.2.2 Results. The special instructions given to the debiasing group had no effect on performance. Neither the debiasing group nor the control group was able to improve consistently upon the mean responses given by subjects in Experiment 1. For each pair, we calculated the percentage of subjects in the debiasing group and in the control group whose responses were closer to the true ratio than was the geometric mean of the original, Experiment 1, group. In every case, the percentage of subjects whose responses were closer to the true ratio was the same as the percentage of subjects whose responses were closer to the ratio predicted from the regression line (i.e., who had smaller secondary bias). The average percentage of improved answers was only 53.8 for the experimental group (range 21% to 82%) and 52.4 for the control group (range 37% to 70%). The experimental group showed a better improvement percentage than the control group on 10 pairs, the control group was better for eight pairs, and there was a tie on one pair.

7.0 DISCUSSION

7.1 Psychological significance

As in previous studies, our subjects exhibited some competence in judging frequency. We found that the perceived frequency of the various causes of death, words and occupations generally increased with their statistical frequency; similarly, the discriminability of causes increased with the ratio of their statistical frequencies. Furthermore, our subjects' assessments of the frequencies of causes of death, both direct estimates and paired comparisons, correlated more highly with the true answers than did any other measures, such as newspaper reportage and ratings of direct experience with the causes of death.

In addition, a strong primary bias, consisting of overestimation of low frequencies and underestimation of both high frequencies and large ratios, was evident, much as has been found before by Attneave (1953), Teigen (1973) and others (Poulton, 1973). Several reasons for this primary bias can be advanced. First, subjects may avoid using extremely high (or low) numbers in making their responses. That the underestimation of high ratios in Experiment 1 was not simply an artifact of averaging correct and incorrect answers, is shown by the persistence of the effect for pairs in which nearly everyone got the correct answer.

Another possible explanation of the primary bias assumes a two-stage process of frequency estimation: subjects first choose some representative value and then adjust upward or downward according to whatever considerations seem

relevant to the case at hand. Studies of anchoring and adjustment procedures have shown that such adjustments tend to be insufficient (Tversky & Kahneman, 1974). A number of studies of frequency estimation can be interpreted as showing a tendency to anchor on the average frequency in the lists learned (see Rowe & Rose, 1977). Insufficient adjustment would produce too flat a curve, a finding often noted in laboratory studies (see Hintzman, 1976). Perhaps the clearest evidence of anchoring may be found in Experiment 3, in which the one true frequency given to the subjects could easily have served as an anchor value. Group MVA, who were given a high anchor (50,000), generally assigned higher values to the items than did Group E, whose anchor value was 1,000.

In the paired-comparison tasks no such clear-cut anchor was provided. Nonetheless, Poulton (1968) has shown that in magnitude estimation studies the subjective magnitude of the first stimulus presented serves as an anchor for subsequent judgments. This view is supported by Carroll's (1971) finding of a .66 correlation between the log of individual subjects' first estimate and the mean log of all their responses in estimating word frequency. The present paired-comparison data are consistent with the notion that the response to the first stimulus serves as an anchor. The two causes-of-death groups perceived the first stimulus (pair 40, true ratio = 5.3) as having a low ratio (the geometric mean response for students was 4.3; for League members, 18.0); these two groups showed more underestimation of high ratios than the words and occupations groups, whose geometric mean responses to the first pair were 116 and 265, respectively.

Yet another possible explanation of the primary bias derives from the availability heuristic (Tversky & Kahneman, 1973), which states that assessments of frequency or probability are based on the number of instances of the event that come to mind. Cohen (1966) has found that when subjects manage to recall any of the words in a category the mean number of words recalled per category is relatively independent of the number of words in that category. If this tendency is true also for categories learned outside the laboratory, such as causes of death, and if, as suggested by Tversky and Kahneman, people base their assessments on these all-too-equal recollections, a flattening of their responses, as observed, would result.

The present findings also demonstrated strong and consistent secondary biases that disrupted the monotonic relationships discussed above. Some portion of these biases may be due to the biased coverage of these causes of death in the news media. Others have also speculated about the effects of such media bias. For example, Zebroski (1975) blamed the media for people's concerns about nuclear reactor safety. He noted that "fear sells"; the media dwell on potential catastrophes and not on the successful day-to-day operations of power plants. Author Richard Bach made a similar observation about the fear shown by a young couple going for their first airplane ride:

> In all that wind and engineblast and earth tilting and going small below us, I watched my Wisconsin lad and his girl, to see them change. Despite their laughter, they had been afraid of the airplane. Their only knowledge of flight came from newspaper headlines, a knowledge of collisions

> and crashes and fatalities. They had never read a single report of a little airplane taking off, flying through the air and landing again safely. They could only believe that this must be possible, in spite of all the newspapers, and on that belief they staked their three dollars and their lives (Bach, 1973, p. 37).

The present results suggest that the media have important effects on our perceptions not only because of what they don't report (successful plane trips or reactor operations), but because of what they do report to a disproportionate extent.

Subjects may also be misinformed because of bias in their direct exposure to the various causes of death. Although direct death was the rating measure most highly correlated with true frequency, those correlations were still well below unity (.62 for paired comparisons, .82 for direct estimates). Young people, such as our student subjects, may be underexposed to death from various diseases associated with age, like stroke, stomach cancer and diabetes, all of which were underestimated, and overexposed to death from motor vehicle accidents, all accidents, and pregnancy, all of which were overestimated relative to the regression line.

The two explanations of secondary bias given above assume that the bias occurs because the information received by the subject is inadequate or misleading. A more psychologically interesting explanation can be found by examining hypotheses about the biases induced by people's cognitive storage and retrieval processes. Tversky and Kahneman's (1973) concept of availability seems relevant here. According to this heuristic, events that are more imaginable, vivid, or sensational are more easily recalled and thus are

relatively overestimated, while drab or unspectacular events are underestimated. Examination of Figures 4-1 and 4-2 supports this view. Among the most overestimated causes of death (relative to the regression line) are botulism, tornado, flood, homicide, motor vehicle accidents, all accidents and cancer. These are all sensational events. Most of the causes of death that were most underestimated (relative to the regression line), asthma, tuberculosis, diabetes, stomach cancer, stroke and heart disease, seem to be undramatic, quiet killers.

Some of the evidence of secondary bias is consistent with previous laboratory findings. One such finding is that more concrete and imaginable words are perceived as less likely than equally frequent abstract words (e.g., Ghatala & Levin, 1976). While we had no direct measure of imaginability, one might assume that catastrophic events and those more heavily reported in the media tend to be more concrete and imaginable. However, all three of these surrogate measures of imaginability (catastrophe, news frequency and news inches) were positively correlated with the residuals (for both paired comparisons and direct estimates). Thus, in this sense, imaginable events tended to be judged more likely, as predicted by availability considerations.

Another difference between the present research and previous studies is found with catastrophic causes of death whose occurrences tend to be massed rather than distributed over time. Laboratory studies (e.g., Rowe & Rose, 1977) have consistently found that massing the occurrences of a word in a learned list tends to decrease its perceived frequency.

Two explanations offered for this effect (Hintzman, 1976) are (a) encoding variability: spaced repetitions are more likely to receive differential coding than massed items; and (b) deficient processing of massed items. In the current experiments, catastrophic (massed) events tended to be overestimated relative to the regression line. The key difference between the usual laboratory experiments and the present study is that the former do not use stimuli that become sensational or emotionally charged when massed. Such special characteristics may lead to extra processing, rather than to deficient processing, for catastrophic causes of death.

When we have been able to compare the present results with previous laboratory work, we have found about as many mismatches as matches. The present study is based on material our subjects have learned in the real world; in most other laboratory work, the subjects were tested on material they had learned in the laboratory. Mandler (1976) has speculated on this difference:

> In terms of presentation of to-be-remembered material, the laboratory experiment fails--in comparison with the real world--with respect to three major problems: Frequency, salience, and context. The laboratory experiments fail with respect to frequency because the typical event that an individual must recall or recognize in everyday life has been encountered anywhere from a few to thousands of times; in the laboratory we look at the few and rarely look at the thousands. Salience must be of interest because encoding operations in the real world typically take place with particular attention to the relevance or salience of a particular event to other aspects of the mental apparatus; we encode what is important, while in the laboratory we are required

to encode what is unimportant. Furthermore, the context of real world memory involves not simply a restricted number of materials presented in the laboratory, together with a computer or a memory drum, but rather the larger context of the individual's current plans and intentions, geographic location, and social conditions (pp. 3-4).

7.2 Improving judgments

One question raised by this study is how to improve intuitive judgments of frequency. We did not attempt here to correct the primary (overestimation/underestimation) bias. Work by Teigen (1973) suggests that this can be done by asking people to allocate frequencies as percentages of the total rather than having them estimate absolute numbers. This technique, however, might not prove helpful when (as with causes of death) the largest frequency is over a million times larger than the smallest frequency. It would be exceedingly difficult for subjects to express ratios even as high as 3000 to 1 (as they did in the present study) using a percentage response mode. As mentioned earlier, statistical correction might be the best way to correct the primary bias.

Since the secondary bias observed here seems linked to availability, we hoped to reduce that bias by informing subjects about its probable source. This information was not useful. The failure of such frontal attacks to eliminate biases (see also Fischhoff, 1977) suggests some directed restructuring of judgment tasks may be necessary. For example, Selvidge (Note 1) proposed having people make probability and frequency judgments on a scale in which other familiar events serve as marker points. In composing such a scale, great care would have to be taken to use only events

whose subjective ordering fits their true ordering. Beyth-Marom and Fischhoff (1977) have shown that requiring people to work hard to produce specific examples of classes of events before estimating the frequencies of the classes can partially reduce availability bias. Another promising suggestion comes from Armstrong, Denniston and Gordon (1975), who found that numerical estimates can be improved by having estimators decompose the original question into a series of sub-questions about which they are more knowledgeable and whose answers lead logically to the estimate of interest. For example, an answer to the question "How many people were killed in motor vehicle accidents in the United States in 1970?" might be improved by having people answer the related questions:

(a) What is the population of the U.S.?
(b) How many automobile trips does the average U.S. citizen take in a year?
(c) What is the probability of a fatal injury on any particular trip?

From the answers to these questions, one can calculate an answer to the original question.

7.3 Societal implications

Economist Frank Knight once observed that "We are so built that what seems reasonable to us is likely to be confirmed by experience or we could not live in the world at all" (Knight, 1921, p. 227). But the present study and a growing body of other research (e.g., Kates, [illegible] Kunreuther et al., 1978; Slovic, Kunreuther & White, 1974) indicate that

in the perception of risks and hazards, Knight's optimistic assessment of human capabilities is wrong. People do not have accurate knowledge of the risks they face. As our society puts more and more effort into the regulation and control of these risks (banning cyclamates in food, lowering highway speed limits, paying for emergency coronary-care equipment, etc.), it becomes increasingly important that these biases be recognized and, if possible, corrected. Improved public education is needed before we can expect the citizenry to make reasonable public-policy decisions about societal risks. And the experts who guide and influence these policies should be aware that when they rely on their own experience, memory and common sense rather than on statistical data, they, too, may not be immune to bias.

8.0 REFERENCE NOTES

1. Selvidge, J. Assigning probabilities to rare events. Unpublished Ph.D. dissertation, Harvard University, 1972.

2. Combs, B. & Slovic, P. Causes of death: Biased newspaper coverage and biased perceptions. Unpublished manuscript, Decision Research, Eugene, Oregon, 1977.

9.0 REFERENCES

Arlen, M. J. The cold, bright charms of immortality. The New Yorker, 1975, 50(49), 73-78.

Armstrong, J. S., Denniston, W. B. & Gordon, M. M. The use of the decomposition principle in making judgments. Organizational Behavior and Human Performance, 1975, 14, 257-263.

Attneave, F. Psychological probability as a function of experienced frequency. Journal of Experimental Psychology, 1953, 46, 81-86.

Bach, R. Nothing by chance. The American Way, 1973, 6, 32-38.

Bar-Hillel, M. The base-rate fallacy in probability judgments. Decision Research Report 77-4, 1977.

Bergstrom, T. C. Preference and choice in matters of life and death. In J. Hirshleifer, T. Bergstrom & E. Rappoport, Applying Cost-Benefit Concepts to Projects Which Alter Human Mortality, Report UCLA-ENG 7478. Los Angeles, California: University of California, School of Engineering and Applied Science, 1974.

Beyth-Marom, R. & Fischhoff, B. Direct measures of availability and frequency judgments. Bulletin of the Psychonomic Society, 1977, 9, 236-238.

Burton, I., Kates, R. W. & White, G. F. The Environment as Hazard. New York: Oxford University Press, in press.

Carroll, J. B. Measurement properties of subjective magnitude estimates of word frequency. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 722-729.

Cohen, B. H. Some-or-none characteristics of coding behavior. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 182-187.

Coombs, C. H., Dawes, R. M. & Tversky, A. Mathematical Psychology: An Elementary Introduction. Englewood Cliffs, N. J.: Prentice-Hall, 1970.

Estes, W. K. The cognitive side of probability learning. Psychological Review, 1976, 83, 37-64.

Fischhoff, B. Perceived informativeness of facts. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 349-358.

Ghatala, E. S. & Levin, J. R. Phenomenal background frequency and the concreteness/imagery effect in verbal discrimination learning. Memory & Cognition, 1976, 4, 302-306.

Hintzman, D. L. Apparent frequency as a function of frequency and the spacing of repetitions. Journal of Experimental Psychology, 1969, 80, 139-145.

Hintzman, D. L. Repetition and memory. In G. H. Bower (Ed.), The Psychology of Learning and Motivation: Advances in Research and Theory. New York: Academic Press, 1976.

Hintzman, D. L. The Psychology of Learning and Memory. San Francisco: W. H. Freeman, in press.

Howell, W. C. Representation of frequency in memory. Psychological Bulletin, 1973, 80, 44-53.

Kates, R. W. Hazard and choice perception in flood plain management. Research Paper No. 78. Chicago: University of Chicago, Department of Geography, 1962.

Kates, R. W. Risk assessment of environmental hazard. SCOPE Report 8, Sussex, England: Wiley, in press.

Knight, F. H. Risk, Uncertainty, and Profit. Boston & New York: Houghton-Mifflin, 1921.

Kunreuther, H., Ginsberg, R., Miller, L., Sagi, P., Slovic, P., Borkin, B. & Katz, N. Disaster Insurance Protection: Public Policy Lessons. New York: Wiley, 1978.

Kučera, H. & Francis, W. N. Computational Analysis of Present-day American English. Providence, R. I.: Brown University Press, 1967.

Mandler, G. Memory research reconsidered: A critical review of traditional methods and distinctions. Center for Human Information Processing Report 64. San Diego: University of California, 1976.

Peterson, C. R. & Beach, L. R. Man as an intuitive statistician. Psychological Bulletin, 1967, 68, 29-46.

Postman, L. Short-term memory and incidental learning. In A. W. Melton (Ed.), Categories of Human Learning. New York: Academic Press, 1964.

Poulton, E. C. The new psychophysics: Six models for magnitude estimation. Psychological Bulletin, 1968, 69, 1-19.

Poulton, E. C. Unwanted range effects from using within-subject experimental designs. Psychological Bulletin, 1973, 80, 113-121.

Rowe, E. J. & Rose, R. J. Effects of orienting task, spacing of repetitions, and list context on judgments of frequency. Memory & Cognition, 1977, 5, 505-512.

Slovic, P., Kunreuther, H. & White, G. Decision processes, rationality and adjustment to natural hazards. In G. F. White (Ed.), Natural Hazards, Local, National and Global. New York: Oxford University Press, 1974.

Teigen, K. H. Number and percentage estimates in sequential tasks. Perceptual and Motor Skills, 1973, 36, 1035-1038.

Thorndike, E. L. & Lorge, I. The Teacher's Word Book of 30,000 Words. New York: Teacher's College Press, Columbia University, 1944.

Tversky, A. & Kahneman, D. Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 1973, 5, 207-232.

Tversky, A. & Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.

Underwood, B. J. Attributes of memory. Psychological Review, 1969, 76, 559-573.

U. S. Bureau of the Census. Census of population: 1970. Occupation by industry. Final Report PC(2)-7C. Washington, D. C.: U. S. Government Printing Office, 1972.

White, G. F. (Ed.), Natural Hazards: Local, National and Global. New York: Oxford University Press, 1974.

Zebroski, E. L. Attainment of balance in risk-benefit perceptions. In D. Okrent (Ed.), Risk-Benefit Methodology and Application: Some Papers Presented at the Engineering Foundation Workshop, Asilomar, California, Report ENG-7598. Los Angeles: University of California, School of Engineering and Applied Science, 1975, 633-644.

CONTRACT DISTRIBUTION LIST
(Unclassified Technical Reports)

Director Advanced Research Projects Agency Attention: Program Management Office 1400 Wilson Boulevard Arlington, Virginia 22209	2 copies
Office of Naval Research Attention: Code 455 800 North Quincy Street Arlington, Virginia 22217	3 copies
Defense Documentation Center Attention: DDC-TC Cameron Station Alexandria, Virginia 22314	12 copies
DCASMA Baltimore Office Attention: Mr. K. Gerasim 300 East Joppa Road Towson, Maryland 21204	1 copy
Director Naval Research Laboratory Attention: Code 2627 Washington, D.C. 20375	6 copies
Office of Naval Research Attention: Code 102IP 800 North Quincy Street Arlington, Virginia 22217	6 copies
Decisions and Designs, Incorporated 8400 Westpark Drive, Suite 600 McLean, Virginia 22101	20 copies

SUPPLEMENTAL DISTRIBUTION LIST
(Unclassified Technical Reports)

Department of Defense

Director of Net Assessment
Office of the Secretary of Defense
Attention: MAJ Robert G. Gough, USAF
The Pentagon, Room 3A930
Washington, DC 20301

Assistant Director (Net Technical Assessment)
Office of the Deputy Director of Defense Research and Engineering (Test and Evaluation)
The Pentagon, Room 3C125
Washington, DC 20301

Assistant Director (Environmental and Life Sciences)
Office of the Deputy Director of Defense Research and Engineering (Research and Advanced Technology)
Attention: COL Henry L. Taylor
The Pentagon, Room 3D129
Washington, DC 20301

Director, Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Director, Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Director, ARPA Regional Office (Europe)
Headquarters, U.S. European Command
APO New York 09128

Director, ARPA Regional Office (Pacific)
Staff CINCPAC, Box 13
APO San Francisco 96610

Dr. Don Hirta
Defense Systems Management School
Building 202
Ft. Belvoir, VA 22060

Chairman, Department of Curriculum Development
National War College
Ft. McNair, 4th and P Streets, SW
Washington, DC 20319

Defense Intelligence School
Attention: Professor Douglas E. Hunter
Washington, DC 20374

Vice Director for Production
Management Office (Special Actions)
Defense Intelligence Agency
Room 1E863, The Pentagon
Washington, DC 20301

Command and Control Technical Center
Defense Communications Agency
Attention: Mr. John D. Hwang
Washington, DC 20301

Department of the Navy

Office of the Chief of Naval Operations (OP-951)
Washington, DC 20450

Office of Naval Research
Assistant Chief for Technology (Code 200)
800 N. Quincy Street
Arlington, VA 22217

Office of Naval Research (Code 230)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Naval Analysis Programs (Code 431)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Operations Research Programs (Code 434)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Information Systems Program (Code 437)
800 North Quincy Street
Arlington, VA 22217

Director, ONR Branch Office
Attention: Dr. Charles Davis
536 South Clark Street
Chicago, IL 60605

Director, ONR Branch Office
Attention: Dr. J. Lester
495 Summer Street
Boston, MA 02210

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

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

Office of Naval Research
Scientific Liaison Group
Attention: Dr. M. Bertin
American Embassy - Room A-407
APO San Francisco 96503

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps (Code RD-1)
Washington, DC 20380

Headquarters, Naval Material Command
(Code 0331)
Attention: Dr. Heber G. Moore
Washington, DC 20360

Head, Human Factors Division
Naval Electronics Laboratory Center
Attention: Mr. Richard Coburn
San Diego, CA 92152

Dean of Research Administration
Naval Postgraduate School
Attention: Patrick C. Parker
Monterey, CA 93940

Superintendent
Naval Postgraduate School
Attention: R. J. Roland, (Code 52R1)
C^3 Curriculum
Monterey, CA 93940

Naval Personnel Research and Development
Center (Code 305)
Attention: LCDR O'Bar
San Diego, CA 92152

Navy Personnel Research and Development
Center
Manned Systems Design (Code 311)
Attention: Dr. Fred Muckler
San Diego, CA 92152

Naval Training Equipment Center
Human Factors Department (Code N215)
Orlando, FL 32813

Naval Training Equipment Center
Training Analysis and Evaluation Group
(Code N-00T)
Attention: Dr. Alfred F. Smode
Orlando, FL 32813

Director, Center for Advanced Research
Naval War College
Attention: Professor C. Lewis
Newport, RI 02840

Naval Research Laboratory
Communications Sciences Division (Code 5403)
Attention: Dr. John Shore
Washington, DC 20375

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

Chief, Intelligence Division
Marine Corps Development Center
Quantico, VA 22134

Department of the Army

Alan H. Curry
Operations and Management Science Division
U.S. Army Institute for Research in Management Information and Computer Science
730 Peachtree St., N.E. (Suite 900)
Atlanta, Georgia 30308

Deputy Under Secretary of the Army
(Operations Research)
The Pentagon, Room 2E621
Washington, DC 20310

Director, Army Library
Army Studies (ASDIRS)
The Pentagon, Room 1A534
Washington, DC 20310

U.S. Army Research Institute
Organizations and Systems Research Laboratory
Attention: Dr. Edgar M. Johnson
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Organizations and Systems Research Laboratory
U.S. Army Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Technical Director, U.S. Army Concepts Analysis Agency
8120 Woodmont Avenue
Bethesda, MD 20014

Director, Strategic Studies Institute
U.S. Army Combat Developments Command
Carlisle Barracks, PA 17013

Commandant, Army Logistics Management Center
Attention: DRXMC-LS-SCAD (ORSA)
Ft. Lee, VA 23801

Department of Engineering
United States Military Academy
Attention: COL A. F. Grum
West Point, NY 10996

Marine Corps Representative
U.S. Army War College
Carlisle Barracks, PA 17013

Chief, Studies and Analysis Office
Headquarters, Army Training and Doctrine Command
Ft. Monroe, VA 23351

Commander, U.S. Army Research Office (Durham)
Box CM, Duke Station
Durham, NC 27706

Department of the Air Force

Assistant for Requirements Development and Acquisition Programs
Office of the Deputy Chief of Staff for Research and Development
The Pentagon, Room 4C331
Washington, DC 20330

Air Force Office of Scientific Research
Life Sciences Directorate
Building 410, Bolling AFB
Washington, DC 20332

Commandant, Air University
Maxwell AFB, AL 36112

Chief, Systems Effectiveness Branch
Human Engineering Division
Attention: Dr. Donald A. Topmiller
Wright-Patterson AFB, OH 45433

Deputy Chief of Staff, Plans, and Operations
Directorate of Concepts (AR/XOCCC)
Attention: Major R. Linhard
The Pentagon, Room 4D 1047
Washington, DC 20330

Director, Advanced Systems Division (AFHRL/AS)
Attention: Dr. Gordon Eckstrand
Wright-Patterson AFB, OH 45433

Commander, Rome Air Development Center
Attention: Mr. John Atkinson
Griffis AFB
Rome, NY 13440

IRD, Rome Air Development Center
Attention: Mr. Frederic A. Dion
Griffis AFB
Rome, NY 13440

HQS Tactical Air Command
Attention: LTCOL David Dianich
Langley AFB, VA 23665

Other Government Agencies

Chief, Strategic Evaluation Center
Central Intelligence Agency
Headquarters, Room 2G24
Washington, DC 20505

Director, Center for the Study of
Intelligence
Central Intelligence Agency
Attention: Mr. Dean Moor
Washington, DC 20505

Mr. Richard Heuer
Methods & Forecasting Division
Office of Regional and Political Analysis
Central Intelligence Agency
Washington, DC 20505

Office of Life Sciences
Headquarters, National Aeronautics and
Space Administration
Attention: Dr. Stanley Deutsch
600 Independence Avenue
Washington, DC 20546

Other Institutions

Department of Psychology
The Johns Hopkins University
Attention: Dr. Alphonse Chapanis
Charles and 34th Streets
Baltimore, MD 21218

Institute for Defense Analyses
Attention: Dr. Jesse Orlansky
400 Army Navy Drive
Arlington, VA 22202

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

Perceptronics, Incorporated
Attention: Dr. Amos Freedy
6271 Variel Avenue
Woodland Hills, CA 91364

Stanford University
Attention: Dr. R. A. Howard
Stanford, CA 94305

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

Department of Psychology
Brunel University
Attention: Dr. Lawrence D. Phillips
Uxbridge, Middlesex UB8 3PH
England

Decision Analysis Group
Stanford Research Institute
Attention: Dr. Miley W. Merkhofer
Menlo Park, CA 94025

Decision Research
1201 Oak Street
Eugene, OR 97401

Department of Psychology
University of Washington
Attention: Dr. Lee Roy Beach
Seattle, WA 98195

Department of Electrical and Computer
Engineering
University of Michigan
Attention: Professor Kan Chen
Ann Arbor, MI 94135

Department of Government and Politics
University of Maryland
Attention: Dr. Davis B. Bobrow
College Park, MD 20747

Department of Psychology
Hebrew University
Attention: Dr. Amos Tversky
Jerusalem, Israel

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: PTR-1042-77-7
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): (6) PERCEIVED FREQUENCY OF RECURRENT EVENTS,

5. TYPE OF REPORT & PERIOD COVERED: (9) Technical Report,

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): (10) Sarah Lichtenstein, Mark Layman, Paul Slovic, Barbara Combs, Baruch Fischhoff,

8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0074, Subcontract 76-030-0714

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Decision Research, A Branch of Perceptronics, 1201 Oak Street, Eugene, Oregon 97401

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: ARPA Order No. 3052

11. CONTROLLING OFFICE NAME AND ADDRESS: Defense Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, Virginia 22217

(11) 12. REPORT DATE: Jul[illegible]77

13. NUMBER OF PAGES: 94

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research, Code 455, 800 No. Quincy Street, Arlington, Virginia 22217

(12) 94p

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):

(15) N0001[illegible]4-76-C-0074, ARPA Order-3052

18. SUPPLEMENTARY NOTES:

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

Probability Assessment
Perceived Frequencies
Rare Events
Lethal Events
Risk Perception
Availability

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

A series of experiments involving over 650 individuals studied how people judge the frequency of various recurrent events. The judgments exhibited a highly consistent but systematically biased subjective scale of frequency. Two kinds of bias were identified: (a) a tendency to overestimate small frequencies and underestimate larger ones, and (b) a tendency to exaggerate the frequency of some specific events and to underestimate the frequency of others, at

cont'd

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

390 664

20. (continued)

any given level of objective frequency. These biases were traced to a number of possible sources including disproportionate exposure, memorability or imaginability of various events. Subjects were unable to correct for these sources of bias when specifically instructed to avoid them. Comparisons with previous laboratory studies are discussed, along with methods for improving frequency judgments.

A

END

9-78

DDC