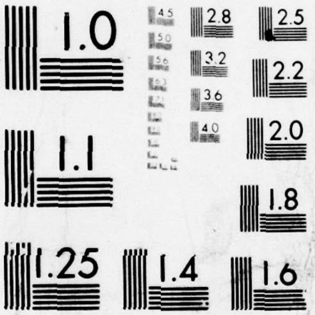
AD-A073 368 DECISIONS AND DESIGNS INC MCLEAN VA F/G 5/10
AN EMPIRICAL INVESTIGATION OF BAYESIAN INFERENCE WITH TWO PEOPL--ETC(U)
DEC 78 W F GABRIELLI, W EDWARDS N00014-76-C-0074

UNCLASSIFIED NL

1 OF 1
AD A073368

END
DATE
FILMED
10-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

C

001922-8-T

USC

UNIVERSITY OF SOUTHERN CALIFORNIA

social science research institute

RESEARCH REPORT

AN EMPIRICAL INVESTIGATION OF BAYESIAN INFERENCE WITH TWO PEOPLE

WILLIAM F. GABRIELLI, JR.
WARD EDWARDS

See 1473

SPONSORED BY:
ADVANCED RESEARCH PROJECTS AGENCY
DEPARTMENT OF DEFENSE

MONITORED BY:
ENGINEERING PSYCHOLOGY PROGRAMS
OFFICE OF NAVAL RESEARCH
CONTRACT NO. N00014-76-C-0074, ARPA

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED;
REPRODUCTION IN WHOLE OR IN PART IS
PERMITTED FOR ANY USE OF THE U.S. GOVERNMENT

DECEMBER 1978

SSRI RESEARCH REPORT 78-8

Social Science Research Institute
University of Southern California
Los Angeles, California 90007
(213) 741-6955

INSTITUTE GOALS:

The goals of the Social Science Research Institute are threefold:

- To provide an environment in which scientists may pursue their own interests in some blend of basic and methodological research in the investigation of major social problems.
- To provide an environment in which graduate students may receive training in research theory, design and methodology through active participation with senior researchers in ongoing research projects.
- To disseminate information to relevant public and social agencies in order to provide decision makers with the tools and ideas necessary to the formulation of public social policy.

HISTORY:

The Social Science Research Institute, University of Southern California, was established in 1972, with a staff of six. In fiscal year 1978-79, it had a staff of over 90 full- and part-time researchers and support personnel. SSRI draws upon most University academic Departments and Schools to make up its research staff, e.g. Industrial and Systems Engineering, the Law School, Psychology, Public Administration, Safety and Systems Management, and others. Senior researches have point appointments and most actively combine research with teaching.

FUNDING:

SSRI Reports directly to the Executive Vice President of USC. It is provided with modest annual basic support for administration, operations, and program development. The major sources of funding support are federal, state, and local funding agencies and private, foundations and organizations. The list of sponsors has recently expanded to include governments outside the United States. Total funding has increased from approximately $150,000 in 1972 to almost $3,000,000 in the fiscal year 1978-1979.

RESEARCH INTERESTS:

Each senior SSRI scientist is encouraged to pursue his or her own reaearch interests, subject to availability of funding. These interests are diverse: a recent count identified 27. Four major interests persist among groups of SSRI researchers: crime control and criminal justice, methods of dispute resolution and alternatives to the courts, use of administration records for demographic and other research purposes, and exploitation of applications of decision analysis to public decision making and program evaluation. But many SSRI projects do not fall into these categories. Most project combine the skills of several scientists, often from different disciplines. As SSRI research personnel change, its interests will change also.

①

An Empirical Investigation of
Bayesian Inference with Two People

Research Report 78-8
December, 1978

William F. Gabrielli, Jr.

Ward Edwards

Social Science Research Institute
University of Southern California

Sponsored by

Defense Advanced Research
Projects Agency

Summary

Twenty subjects from the University of Southern California performed a Bayesian inference task in pairs. Like earlier research in inference, the individuals were asked to infer posterior odds about a pair of hypotheses from a collection of data. Unlike the earlier studies, the individuals were then required to aggregate their posterior odds with those of another individual who had seen a second set of independent data samples to form an opinion about the same pair of hypotheses.

Conservatism and radicalism findings of earlier studies were reconfirmed. Individual subjects' responses collected before aggregation showed conservatism in the high d' condition and radicalism in the low d' condition.

The aggregated final odds from the pairs of subjects seem to reflect some confusion. Some of the subjects apparently used a simple and incorrect averaging strategy. Others did not use this strategy, but in general, pairs of subjects were unable to provide anything but conservative final odds when they aggregated their two opinions.

The importance of using real stimuli, the way the responses were elicited, and the instructions that were given to the subjects are discussed. Also, a "mean of means" or arithmetic log likelihood response mode is discussed as an alternative elicitation mode that may be useful in information aggregation when more than one person is involved

TABLE OF CONTENTS

TABLES

ACKNOWLEDGEMENTS

This research was supported by the Advanced Research Projects Agency of the United States Department of Defense and was monitored by the Office of Naval Research under contract #ONR N00014-76-C-0074 under subcontract P.O. #78-072-0720 from Decisions and Designs, Inc.

The authors would like to thank William Stillwell for his especially valuable help in the development of this study. We would also like to thank Richard John for his comments and criticisms.

DISCLAIMER

INTRODUCTION

The study reported here extends Bayesian inference to the case in which two people rather than one must interpret information and reach a group conclusion. Each of the two people has conditionally independent information relevant to which of the competing hypotheses is true. The information of one subject is different from that of the other subject, and even the combination of data available to both cannot provide certainty. One of the two subjects has more diagnostic information than the other. No conflict exists over what is known. The two must combine their information to make a joint inference about the probabilities of the two hypotheses.

Studies of inference in the one person case have been summarized in Edwards (1968), Anderson (1971) and Slovic and Lichtenstein (1971). Generally, people are conservative as individuals in their judgments. They fail to combine the information to modify prior probabilities as strongly as the data would justify (Wheeler and Edwards, 1975).

Eils, Seaver and Edwards (1977) showed that subjects avoid conservatism if asked to assess an average impact of information, rather than cumulating information over successive data. A cumulative response requires a response number outside the range of the inputs, while an average requires a response somewhere in the middle of that range. Apparently, averaging is far easier than cumulating.

This experiment explores a kind of information aggregation problem for which cumulating is essential. It provides one set of data to one subject and another set, of different diagnosticity, to another subject, and then requires a joint assessment of certainty based on both sets of data. No averaging response mode that makes easy intuitive sense can be designed for this task.

I. FORMAL RULES

Bayes Theorem, the formula for calculating the exact probability of one hypothesis as compared with another given n data and a prior odds, is:

$$\Omega_n = \Omega_o \prod_{i=1}^{n} L_i \qquad (1)$$

where Ω_o is the prior odds in favor of the hypothesis A over its alternative hypothesis not A. L_i is the likelihood ratio for the ith datum, and Ω_n the revised posterior odds.

With log transformation, this formula becomes:

$$\mathrm{Log}\ \Omega_n = \log \Omega_o + \sum_{n=1}^{n} \log L_i \qquad (2)$$

If the subjects are able to provide arithmetic mean log likelihood ratios (AMLL), the posterior odds will be:

$$\Omega_n = \Omega_o 10^{n(\mathrm{AMLL})} \qquad (3)$$

where

$$\mathrm{AMLL} = (1/n) \sum_{i=1}^{n} \log L_i \qquad (4)$$

Here, the final posterior odds ratio is calculated after the subject gives his AMLL estimates.

When each of two people have independent information relevant to a comparison between two hypotheses, the aggregation of that information should be similar to the case in which the same person has two pieces of independent information. If the two people express their feelings based on their own information in likelihood ratios, then they should combine that information using Equation (1), assuming that the information available to each is conditionally independent of that available to the other. If their assessments are expressed as posterior odds (or some quantity from which posterior odds can be infered), the arithmetic is slightly more complicated. First, a prior odds or log prior odds must be known for each; in experiments, this is usually supplied in instructions. Then, either the aggregate likelihood ratio, its logarithm, or the AMLL should be recovered separately for each, using Equation 1, 2, or 3 as appropriate.

Finally, the two assessments should be combined. One way to do so would be to perform the appropriate aggregation for one subject and then use his posterior odds as the prior odds for the other subject. (Since the two subjects receive independent information, this can be done in either order.) Another way is to use Equation 1 or 2 or 3, as appropriate, to obtain an aggregate likelihood ratio for each subject, representing the total aggregate impact of the data that the subject has seen.(Note that his aggregate cannot be an AMLL; it must be a cumulative rather than a mean quantity.) Then the two measures of aggregate impact for the two subjects can be combined, by adding them together if they are in logarithmetic form or by multiplying them together if they are in non-log form. Finally, the output of this combination process across subjects can be combined with the original prior odds.

II. METHODS

Subjects

Twenty undergraduate psychology students at the University of Southern California voluntarily participated as subjects.

All of the subjects received generous participation credit in their introductory psychology class.

Procedure

Subjects participated in pairs. The experimenter told each pair that he had prepared samples from two book bags, each filled with 1,000 poker chips. One of the two book bags contained a majority of blue chips while the other was predominately red. Each subject, working alone, then worked through a response booklet. The first page explained the task. Each subsequent page presented a sample (with replacement) from a book bag; the subjects understood that each successive sample represented a new selection of which bag was being sampled. Every such selection of a bag resulted from a flip of a fair coin.

One of the two subjects in each pair had samples drawn from book bags that had either 750 red chips and 250 blue ones or 250 red chips and 750 blue ones. These data were relatively highly diagnostic; d' = 1.15. The book bags from which the other person's samples were drawn had a 600 to 400 ratio of red chips to blue ones or the reverse. This person's samples were, therefore, relatively less diagnostic given the same number of chips and sample composition (d' = .41). Subjects were informed of the ratio (750:250 or 600:400) from which their samples came. The actual samples the individuals saw are reprinted in Table 1.

The likelihood ratios in Table 1 are easier to verify if one exploits a useful property of symmetric binomial examples like this one. (Symmetric binomial simply means that the two hypotheses are equidistant from 0.5 in opposite directions.) In such cases only, the likelihood ratio for any set of observations is given by the equation

$$L = \left(p/q \right)^{(s - f)} \qquad (5)$$

The symbols p and q refer to the probability of the more and of the less common chip in the sample, respectively. So p/q = 3 for the high diagnosticity bag and 1.5 for the low diagnosticity bag. The symbols s and f stand for successes and failures (time-honored terms from statistical applications of binominal arithmetic);

Table 1. Actual 20 Samples for the 10 Pairs of Subjects

Sample	High Diagnosticity	True Likelihood Ratio	Low Diagnosticity	True Likelihood Ratio
1	BRBR	1.00 (N)	BBBR	3.375 (B)
2	RRBRB	3.00 (R)	RBRRRRRR	11.39 (R)
3	RBBRBBBB	81.00 (B)	BRRBB	1.50 (B)
4	RRBRR	27.00 (R)	RRBB	1.00 (N)
5	RRRRR	243.00 (R)	BRBRRRR	3.375 (R)
6	RRRR	81.00 (R)	RRBRRB	2.25 (R)
7	RRRB	9.00 (R)	BBRR	1.00 (N)
8	BRBRBBB	27.00 (B)	BBBBB	7.59 (B)
9	RBBRRR	9.00 (R)	RRBRRB	2.25 (R)
10	BRBRR	3.00 (R)	RBRB	1.00 (N)
11	RBRRB	3.00 (B)	RRRBBB	1.00 (N)
12	BBBBRBBB	729.00 (B)	RBRBRBB	1.50 (B)
13	BRBRRB	1.00 (N)	RRRR	5.06 (R)
14	RRRBRRBR	81.00 (R)	RRRRBBRR	5.06 (R)
15	BRRBR	3.00 (R)	RBBBBRRR	1.00 (N)
16	RRRR	81.00 (R)	BBBRRB	2.25 (B)
17	BBBBBB	729.00 (B)	RRBBRBB	1.50 (B)
18	BBRBB	27.00 (B)	BRBRB	1.50 (B)
19	RBBRB	3.00 (B)	RBRRRR	5.06 (R)
20	BRRRRBRR	81.00 (R)	RRBRBBRR	2.25 (R)

Note: The symbol R means that a Red chip was drawn; B means that a Blue chip was drawn. Order is irrelevant. The high diagnosticity samples come from the 750-250 bag; the low diagnosticity samples come from the 600-400 bag. (R) following a likelihood ratio means that it favors the predominantly Red bag; (B) means that it favors the predominantly Blue bag, and (N) means that the sample is neutral, i.e. it contains equal numbers of chips of each color.

in this instance, a success is an occurrence of the event more common in the sample, while a failure is an occurrence of the less common event in the sample. Obviously a sample with more reds than blues favors the predominantly red bag and a sample with more blues than reds favors the predominantly blue bag.

After each subject individually had recorded his responses to all 20 samples, the two subjects in each pair were brought together. The experimenter explained that each sample shown to one subject corresponded to the same numbered sample shown to the other subjects, in the sense that both had come from book bags having the same predominant color.

This information, given to ideal Bayesians, would imply that they could combine the information in each pair of samples just by multiplying the two likelihood ratios together. The resulting product, multiplied by the prior odds, would yield an appropriate joint posterior odds.

For example, if a subject judged that sample number two was twice as likely to have come from the predominantly red bag, as from the predominantly blue bag, and his partner judged a likelihood ratio in favor of the same bag of 25:1, the group likelihood ratio should have been 50:1 in favor of that hypothesis, just as if only one individual had obtained both pieces of information.

The experimenter then seated both subjects at a table, each with his previously filled in response booklet in front of him. Yet another response booklet was provided, and the experimenter asked the pair to reach an agreed-on assessment of the probability that the predominant color of the book bags represented by the pair of samples was (say) red. In arriving at this agreed-on assessment, neither subject was permitted to report the sample he had based his judgments on to the other subject. But they were permitted to report their assessments of likelihood ratios based on each sample. They were instructed to consider the two samples conditionally independent, and to consider their individually assessed likelihood ratios as correct, for the purpose of reaching a group assessment.

III. RESULTS

Four different sets of numbers can be compared with one another for each pair of subjects and each sample. They are:

A. Individual responses to each sample.

B. The group's actual response to each sample.

C. A number calculated by multiplying together the likelihood ratios estimated by each of the two members of the group for the particular sample.

D. A correct Bayesian solution, obtained by using Bayes's Theorem for each sample. In the individual response comparisons, this produces the appropriate likelihood ratio. In the group comparisons, it is also necessary to multiply together the likelihood ratios thus obtained for each member of a pair.

These numbers were compared with one another via regression analyses, keeping individual subjects or pairs of subjects distinct but aggregating over samples. As might be expected of untrained subjects, performing an abstract task with little motivation to do it carefully, the data are highly variable. Nevertheless, certain patterns emerge from them that can be most easily displayed by looking at medians of various kinds. Table 2 represents such medians, spelling out which set of numbers specified above is being correlated with or regressed on what other set.

TABLE 2. Median Correlations & Regressions

Variables being related	r	r^2	Intercept	Slope
Individual responses and correct Bayesian solution, high diagnosticity subjects only (A with D)	.59	.35	.69	.31
Individual responses and correct Bayesian solution, low diagnosticity subjects only (A with D)	.76	.58	.21	1.49
Actual group responses and products of individually estimated likelihood ratios (B with C)	.79	.61	.16	.45
Actual group responses and correct Bayesian solution (B with D)	.54	.29	.48	.23
Products of individually estimated likelihood ratios and correct Bayesian solution (C with D)	.53	.28	.79	.50

Note: In calculating regressions, the first quantity listed is the predicted variable and the second is the x-axis variable. All calculations are based on the logs of the specified quantities, expressed as numbers

The median correlations show modest to medium relations between the variables, as is to be expected from highly variable data. (Note that they would have been much higher if they had not been folded at 0.) The slopes for individual responses show a familiar pattern, standard for virtually every experiment of this kind that has ever been performed. For high diagnosticity data, the subjects were conservative (regression slope less than 1). For low diagnosticity data, the subjects were excessive. Other examples of the same finding include Edwards (1968), and Wheeler and Edwards (1975).

The fact that actual group responses correlate fairly highly with products of individually estimated likelihood ratios means no more than that the two members of each group were allowed to tell each other what their individual estimates had been, and were instructed to take those estimates as veridical for the purpose of arriving at group assessments. But the low slope shows that they did not arrive at these final assessments by the normatively appropriate procedure. Direct observation showed that some pairs, but not all, arrived at final estimates by simply averaging their individual estimates--an easy, incorrect approach. To check how effective a theory simple averaging might be for explaining the group assessments, we also calculated regressions between actual group responses (y) and the mean of the logs, rather than the sum of the logs, of the individual responses (x). The median slope for that regression is .89. Similarly, we calculated the same regression between actual group responses and correct Bayesian numbers; that median is .47. These regressions are too low to permit the conclusion that subjects systematically averaged their estimates. We are left with the conclusion that the subjects found the task confusing, and adopted ill-specified and confused strategies for determining their responses.

The y-intercepts of the regression lines are interesting numbers. For virtually all individual and group responses, they are positive. This fact, combined with the regression slopes uniformly less than 1, implies that the regression lines cross the normatively correct 45° line. Why? Inspection of the individual scatterplots, while emphasizing the disorderliness of the data, suggest that this is not a statistical artifact resulting from an attempt to fit a straight line to non-linear data as some versions of the response bias theory of conservatism (See Edwards 1968) might suggest.

An interesting point, not previously discussed in the man-vs.-Bayes literature, is that response modes, like data analyses, can be folded or unfolded. A folded response mode, like that used in this study, requires the subject first to specify the favored hypothesis, and thereafter to specify some appropriate number saying how strongly the evidence favors it. An unfolded response mode does not first require commitment to either hypothesis An interesting hypothesis, consistent with the findings of initial overconfidence by Tversky and Kahneman (1974), is that the initial task of specifying the favored hypothesis drives up the assessment of how likely the hypothesis is. Thereafter, revisions of opinion based on new data proceed conservatively. This idea would be easy to test, but has not been tested.

V. CONCLUSIONS

The main function of this experiment was to explore whether subjects could properly aggregate numbers representing individual degrees of certainty in order to obtain a group number representing the result of combining information inside several heads. They could not, with the response modes used in this study. Instead, they attemped to reach some sort of compromise based on their individual responses.

The study reconfirmed old findings concerning the relation between diagnosticity and conservatism: high diagnosticity leads to conservatism, whereas low diagnosticity leads to radicalism.

The variability and generally poor quality of the data show the importance of careful instruction, motivation and feedback, and the desirability of using real stimuli (e.g actual book bags and poker chips) rather than a paper-and-pencil task designed to have the appropriate formal characteristics. This point has also been made before; see Slovic, Lichtenstein and Edwards, 1965.

It would be premature to conclude from these data that subjects cannot aggregate evidence properly. It would depend on the response mode. Eils, Seaver, and Edwards (1977) found that if subjects were asked to estimate arithmetic mean log likelihood ratios (Equation 4) they could do a quite good job of aggregating normally distributed evidence; it is likely that the same holds true for stimuli in symmetric binomial experiments like this one. It is natural to extend that response mode to a situation in which evidence must be aggregated in several heads rather than one. If the subjects in this experiment had been asked to estimate mean log likelihood ratios, then they could have done an excellent job of assessing the mean of a group of their individual means based on multi-chip samples. Simple arithmetic, performed on those means of means, combined with knowledge of the number of data on which each was based, would then give an approximation to an appropriate aggregate likelihood ratio. How good the approximation is depends on how far the data deviate from 1:1 likelihood ratio. Formally, the appropriate calculation would be for each subject separately to multiply his mean log likelihood ratio by the number of observations on which it was based, and then for the pair of subjects to add these products together. The result would be an appropriate aggregate log likelihood ratio; added to the log prior odds, it would produce the correct log posterior odds. If the subjects instead average their mean log likelihood ratios, and either they or the experimenter then multiply this mean of means by the total number of observations that the two subjects together had made, a too-high approximation will result. How much too high it will be depends on how much each individual subject's mean log likelihood ratio differs from 0. The farther away both are, the nearer the approximation will be to the correct number.

REFERENCES

Anderson, N.H. Averaging versus adding as a stimulus-combination rule in impression formation. Journal of Experimental Psychology, 1965, 70, 394-400.

Edwards, W. Conservatism in human information processing. In:. B Kleinmuntz (ed.), Formal Representation of Human Judgments, New York, Wiley Press, 1968.

Eils, L., Seaver, D. and Edwards, W. Developing the technology of probabilistic inference: Aggregating by averaging reduces conservatism. University of Southern California, Social Science Research Institute Technical Report 77-3. August, 1977.

Slovic, P. Lichtenstein, S.C. and Edwards, W. Boredom-induced changes in preferences among bets. American Journal of Psychology, 1965, 78, 208-217.

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

Stillwell, W.G., Seaver, D. and Edwards, W. The effects of response scales on likelihood ratio judgments. University of Southern California, Social Science Research Institute, Technical Report 77-5, August, 1977.

Tversky, A. and Kahneman, D. Judgment under uncertainty: Huristics and biases. Science, 1974, 185, 1124.

Wheeler, G.E. and Edwards, W. Misaggregation explains conservative inference about normally distributed populations. University of Southern California, Social Science Research Institute, Technical Report 75-11, August, 1975.

CONTRACT DISTRIBUTION LIST
(Unclassified Technical Reports)

Director — 2 copies
Advanced Research Projects Agency
Attention: Program Management Office
1400 Wilson Boulevard
Arlington, Virginia 22209

Office of Naval Research — 3 copies
Attention: Code 455
800 North Quincy Street
Arlington, Virginia 22217

Defense Documentation Center — 12 copies
Attention: DDC-TC
Cameron Station
Alexandria, Virginia 22314

DCASMA Baltimore Office — 1 copy
Attention: Mr. K. Gerasim
300 East Joppa Road
Towson, Maryland 21204

Director — 6 copies
Naval Research Laboratory
Attention: Code 2627
Washington, D.C. 20375

Revised August 1978

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

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
Camp H. M. Smith, Hawaii 96861

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

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

Superintendent
Naval Postgraduate School
Attention: R. J. Roland, (Code 52Rl)
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 5-
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

Deputy Under Secretary of the Army
(Operations Research)
The Pentagon, Room 2E611
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

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

Professor Raymond Tanter
Political Science Department
The University of Michigan
Ann Arbor, MI 48109

Professor Howard Raiffa
Morgan 302
Harvard Business School
Harvard University
Cambridge, MA 02163

Department of Psychology
University of Oklahoma
Attention: Dr. Charles Gettys
455 West Lindsey
Dale Hall Tower
Norman, OK 73069

Institute of Behavioral Science #3
University of Colorado
Attention: Dr. Kenneth Hammond
Room 201
Boulder, Colorado 80309

Decisions and Designs, Incorporated
Suite 600, 8400 Westpark Drive
P.O. Box 907
McLean, Virginia 22101

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER
001922-8

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
(6) An Empirical Investigation of Bayesian Inference with Two People,

5. TYPE OF REPORT & PERIOD COVERED
Technical
10/77 - 12/78

6. PERFORMING ORG. REPORT NUMBER
78-8

7. AUTHOR(s)
(10) William Gabrielli, Jr.
Ward Edwards

8. CONTRACT OR GRANT NUMBER(s)
(15) ~~Prime Contract~~
N00014-76-C-0074

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Social Science Research Institute
University of Southern California
Los Angeles, California 90007

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

12. ~~REPORT DATE~~
(11) Dece[redacted]978

13. NUMBER OF PAGES
12

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
Decisions & Designs, Inc.
8400 Westpark Drive, Suite 600
McLean, VA 22101

15. SECURITY CLASS. (of this report)
unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited

(9) Research rept. Oct 77 - Dec 78,

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(12) 28p.

18. SUPPLEMENTARY NOTES

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

conservatism
likelihood ratio
book bags
high diagnosticity bag
low diagnosticity bag

390664

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

Twenty subjects from the University of Southern California performed a Bayesian inference task in pairs. Like earlier research in inference, the individuals were asked to infer posterior odds about a pair of hypotheses from a collection of data. Unlike the earlier studies, the individuals were then required to aggregate their posterior odds with those of another individual who had seen a second set of independent data samples to form an opinion about the same pair of hypotheses. Conservatism and radicalism

over

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

unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

findings of earlier studies were reconfirmed. Individual subjects responses collected before aggregation showed conservatism in the high d' condition and radicalism in the low d' condition. The aggregated final odds from the pairs of subjects seem to reflect some confusion. Some of the subjects apparently used a simple and incorrect averaging strategy. Others did not use this strategy but in general, pairs of subjects were unable to provide anything but conservative final odds when they aggregated their two opinions. The importance of using real stimuli, the way the responses were elicited, and the instructions that were given to the subjects are discussed. Also, a "mean of means" or arithmetic log likelihood response mode is discussed as an alternative elicitation mode that may be useful in information aggregation when more than one person is involved.