JOURNAL OF MATHEMATICAL PSYCHOLOGY: 2, 180-189 (1965)

Decisions Based on Conflicting and Inaccurate Observations[1]

WAYNE A. WICKELGREN[2]

University of California, Berkeley, California

AND

GORDON M. BECKER[3]

Boston University, Boston, Massachusetts

Subjects learned the accuracies of 8 cues in a series of 50 learning trials and then used pairs of these cues to predict which of two equally likely symbols occurred in each of 100 test trials. It was concluded that: (a) estimates of cues with very high or very low accuracies are better than estimates of cues with intermediate accuracies; (b) accurate cues are perceived more realistically than are inaccurate cues; (c) people tend to maximize expected payoff when faced with conflicting information in binary choice problems; (d) conformity pressures, i.e., the desire to agree with two reports, strongly interfere with maximization if there is uncertainty about the maximizing response, but conformity pressures exert little influence when there is no uncertainty about the maximizing response.

Real life problems often require decisions before all the relevant information is gathered or reviewed. The information that is obtained often comes from a variety of sources, each of which provides inaccurate and incomplete cues. The experiment reported in this paper is an investigation of the way people use combinations of inaccurate sources in a simple choice situation.

Information from a highly untrustworthy source is ordinarily nearly useless. But if the source is untrustworthy enough and the number of possibilities small enough, such information may be useful exactly because of its untrustworthiness. The largest effect of this kind arises when the world must be in one or the other of two states.

[1] This investigation was supported in part by Research Grant M-2348 (RI) from the National Institute of Mental Health, Public Health Service, and in part by the Cowles Foundation for Economic Research, Yale University.

[2] Present address: Department of Psychology, Massachusetts Institute of Technology, Cambridge, Massachusetts.

[3] Present address: General Electric, Corporate Planning Operation, New York, New York.

A girl who *always* says no when she means yes, and vice versa, gives just as complete guidance to a suitor who knows that fact about her as does a girl who always says what she means. In such binary cases, a scale of trustworthiness is simply related to the probability of error, provided that the two kinds of errors are equally likely, as they are throughout the experiment to be reported. Trustworthiness is least when the probability of error is 0.5 and increases linearly with the distance of that probability from 0.5 in either direction; 0.27 and 0.73 probabilities of error would characterize equally trustworthy cues. Of course, a decision maker trying to use the cue with the 0.73 probability of error would have to remember that he should reverse the cue, that is, interpret it to mean the opposite of what it says—an extra information-processing step which might make such cues harder to use than those with probabilities of error less than 0.5.

Suppose there are two independent information sources, both with known fallibility. What should a subject do in order to maximize his probability of correct decision? He should ignore completely the cue with probability of error nearer 0.5, and should do whatever the other cue (after reversal, if its probability of error is greater than 0.5) indicates is appropriate. If the implications of the cues agree and both are equally fallible, of course, the best decision is the one they agree on. If the implications of two equally fallible cues disagree, it makes no difference which decision is made. Much more complex rules are necessary for dealing with conflicting information in general (Becker, Stocklin, Vegh-Villegas, & Wickelgren, 1963; Van Meter & Middleton, 1963), but the ubiquity of serious discrepancies between what men do and what they should do suggested that it might nevertheless be appropriate to look for such discrepancies in a simple situation involving conflicting items of information.

The fallibility or accuracy of a source can be described by the probability that its report is true. If $P_i(t_i \mid t)$ is the probability that source i will report t_i when t in fact occurs and if t occurs with probability $P(t)$, then the accuracy of source i is:

$$A_i = \sum_t P(t)\, P_i(t_i = t \mid t). \tag{1}$$

In general, the accuracy of a source depends upon both the relative frequency of the true "states" t and the relative frequencies of the different cue reports t_i. If the conditional probability $P_i(t_i = t \mid t)$ is constant for all t then the accuracy of the source is:

$$A_i = P_i(t_i = t \mid t) \qquad \text{for any} \qquad t. \tag{2}$$

In the special case described by (2), the accuracy of the source can be estimated from either the proportion of times the cue correctly predicts the state or from the proportion of times the world state correctly predicts the cue report.

If there are but two independent information sources and but two equally likely world states, and if the payoff to the decision maker depends only upon whether or not his prediction of the true state is correct, the decision maker can maximize his payoff by considering only that source with accuracy most deviant from 0.5, provided the accuracy of each source conforms to Eq. 2. This can be seen from the following considerations.

Let the payoff to the decision maker be $O(t, d)$ when the true state is t and decision d is made ($t = 1, 2$; $d = 1, 2$), and let the pay-off $O(t, d)$ depend only upon whether or not $t = d$; i.e. $O(1, 1) = O(2, 2) > O(1, 2) = O(2, 1)$. The expected payoff for decision d upon receipt of report t_1 from source 1 and report t_2 from source 2 is:

$$E(d \mid t_1, t_2) = \sum_t P(t)\, P(t_1, t_2 \mid t)\, 0(t, d). \tag{3}$$

If the information sources are independent, it follows that:

$$P(t_1, t_2 \mid t) = \prod_i P_i(t_i \mid t). \tag{4}$$

In order to maximize expected payoff, the decision maker should choose that d which maximizes (3). If (2) and (4) are true, a comparison of expected payoffs reduces to a comparison of the following:

$$\frac{P_1(t_1 \mid t = 1)\, P_2(t_2 \mid t = 1)}{P_1(t_1 \mid t = 2)\, P_2(t_2 \mid t = 2)} [0(1, 1) \,..\, 0(1, 2)] \tag{5}$$

and

$$\frac{P(t = 2)}{P(t = 1)} [0(2, 2) - 0(2, 1)]. \tag{6}$$

Whenever Eq. 5 > Eq. 6 setting $d = 1$ will maximize (3) and when Eq. 5 < Eq. 6 setting $d = 2$ will maximize (3). If Eq. 5 = Eq. 6, then $E(d_1 \mid t_1, t_2) = E(d_2 \mid t_1, t_2)$ and either $d = 1$ or $d = 2$ is appropriate. Thus, the decision maker need only compare (5) and (6) to decide whether $d = 1$ or $d = 2$ maximizes his payoff. When $P(t = 1) = P(t = 2)$ and $O(11) = O(22) > O(1, 2) = O(2, 1)$, the consideration of whether or not Eq. 5 is greater or less than Eq. 6 is equivalent to a consideration of whether or not the following is greater or less than unity.

$$\frac{P_1(t_1 \mid t = 1)\, P_2(t_2 \mid t = 1)}{P_1(t_1 \mid t = 2)\, P_2(t_2 \mid t = 2)}. \tag{7}$$

When $t_i = t$ then $P_i(t_i \mid t) = A_i$ and when $t_i \neq t$ then $P_i(t_i \mid t) = 1 - A_i$. Substituting A_i and $1 - A_i$ in (7) results in the following rules for maximizing expected payoff.

When $t_i \neq t_j$

$$\begin{cases} d = t_i & \text{if} \quad A_i > A_j \\ d = t_j & \text{if} \quad A_j > A_i \end{cases} \tag{8}$$

When $t_i = t_j$

$$\begin{cases} d = t_i & \text{if} \quad A_i + A_j > 1 \\ d \neq t_i & \text{if} \quad A_i + A_j < 1 \end{cases} \tag{9}$$

Thus if t_H and t_L are the reports of the more accurate and less accurate sources, respectively, and A_H and A_L are the accuracies of the two sources, the maximizing decision is to disagree with t_H, the report of the more accurate source, only if the more accurate source both reports the same state as the less accurate ($t_H = t_L$) and (a) $A_H \leqslant 0.5$ or (b) $0 < (A_H - 0.5) < (0.5 - A_L) > 0$. Agreement with the more accurate report will maximize expected payoff under all other conditions. There are therefore only the four cases shown in Table 1 that need be considered.

TABLE 1

MAXIMIZING DECISIONS AS A FUNCTION OF CUE ACCURACIES AND REPORTS

Case	Reports	Accuracies	Maximizing decision
I	$t_i \neq t_j$	$A_j \leqslant A_i$	$d = t_i$
II	$t_i = t_j$	$0 \leqslant (A_i - 0.5) \geqslant \mid A_j - 0.5 \mid$	$d = t_i$
III	$t_i = t_j$	$A_j \leqslant A_i \leqslant 0.5$	$d \neq t_i$
IV	$t_i = t$	$0 \leqslant (A_i - 0.5) \leqslant (0.5 - A_j) \geqslant 0$	$d \neq t_i$

Earlier studies by Bruner, Goodnow, & Austin (1956) and by Cohen, Wickelgren & Becker (1963) investigated situations in which both accuracies were greater than 0.5. It was found that people in these situations did maximize expected payoff by agreeing with the report of the more accurate source. When the accuracies were 0.5 for both sources and the reports disagreed, an "all-or-none" type of behavior was found; i.e., subjects tended to choose one of the sources and always agree with its report. The present experiment extends the investigations of Bruner *et al.* and Cohen *et al.* by including accuracies less than 0.5, and by including situations in which an agreeing attitude leads to nonmaximizing behavior.

METHOD

Subjects. Six male summer students at Yale University were obtained through the part-time student employment office of the University. The students were hired as subjects in an experimental study by the economics department[4] and were promised \$1.25 per hour for at least ten hours. After being hired, they were given the opportunity to earn up to \$2 per hour in bonuses depending upon how well they did as compared to other subjects in the experiment. None of the subjects was majoring in psychology, statistics, or economics.

Task. The subjects in the experiment were told that a + or a − symbol would be drawn from a population of + and − symbols. On each trial they were to guess whether a + or a − would occur. After indicating which symbol they expected on the current trial, the subject was shown the "reports" of two or more "observers." Each observer reported whether a + or a − had occurred on that trial. The observers were distinguished by the color of the symbol they reported, e.g., a blue + represented a report by the blue observer that a + had occurred on that trial. All the observers did not always report the true symbol. The accuracies of their reports, or cues, were 1.0, 0.8, 0.6, 0.5, 0.4, 0.2, and 0 for the blue, red, orange, purple, black, brown, and blue-green cues respectively, i.e., the blue symbol was always the symbol that actually occurred on that trial, the blue-green symbol was never the symbol that actually occurred, and the purple symbol was the true symbol on half of the trials. The eighth cue was pink and had an accuracy of 0.5 (same as purple).

After the subject was shown the cue reports he indicated which of the symbols + or − he believed had actually occurred. The bonus to the subject depended upon the number of trials on which his second guess was correct.

Procedure. Each subject was tested individually. The first 50 trials were designed to permit the subject to learn the accuracies of the eight cues. On each trial, after the initial guess by the subject, he was shown the reports of all eight observers. The reports were printed on a single 3 × 5 card, four in one row and four in another. The subject estimated the accuracy of each of the observers and indicated which of the two symbols he now believed had occurred. After this guess, a second 3 × 5 card indicating the true symbol for that trial was placed beside the observer report card. Both cards were then removed for the next trial.

After the 50 learning trials, the subject made a final estimate of the accuracy of each of the eight observers. Since he would not be shown the true symbol on any of the remaining trials, he was told that he would have no basis for changing his estimates of the reliability of any observer. On each of the remaining trials, he was shown only two observer reports. In order to be sure that the subject recognized the observers on each trial, he specified the accuracy of each of the two cues the experimenter had selected for that trial. His specification of an accuracy was in terms of the percentage of time reports in that color were correct. If the subject specified an accuracy different from that which he gave to the cue on the 50th learning trial, the experimenter corrected him and reminded him of his former estimate. Thus, on each of the 100 test trials, the subject first guessed which of the two symbols would be drawn, then saw the reports of two observers, indicated the accuracies of each, and indicated which symbol he then believed had been drawn. He was not shown the true symbol for that trial and his identification of an accuracy was corrected if it differed from his last estimate on the learning trials.

Experimental Design. During the 50 learning trials, a 50:50 random plus-minus distribution was used to generate the true sequence. Subjects were permitted to count the number of times each symbol agreed with the true symbol and to use pencil and paper, or other aids if they desired. Aids were not suggested, encouraged, or discouraged.

[4] The assistance of Harold Watts of the Cowles Foundation for Economic Research, Yale University, in providing research facilities and assistance is gratefully acknowledged.

Each of the 100 test trials consisted of showing the subject a 3 × 5 card with only two cue reports. One cue was always the purple cue with 0.5 accuracy. The second cue changed from one trial to the next. Forty of the trials involved conflicting reports in which the second cue had 0.5 or higher accuracy. Sixty trials employed a second cue with less than 0.5 accuracy. Thirty of these 60 trials involved conflicting reports, i.e., the reports differed but the implications were the same; and the other 30 trials involved nonconflicting reports, i.e., the reports were the same but the implications differed. The cues and maximizing decisions are shown in Table 2.

The order of cue pair presentation was random for the first 50 test trials, but every block of ten trials included all of the cue-pairs shown in Table 2. The order of presentation in the second 50 test trials was formed by reversing, in blocks of ten, the order established in the first 50 trials. The true symbol sequence over the entire 100 trials was generated from a 50:50 random plus-minus distribution.

TABLE 2

PLANNED EXPERIMENTAL CONDITIONS

Accuracies		Color[a]	Reports	Maximizing	Case
A_i	A_j	A_j		decision	
0.5	1.0	Blue	$t_i \neq t_j$	$d = t_j$	I
0.5	0.8	Red	$t_i \neq t_j$	$d = t_j$	I
0.5	0.6	Orange	$t_i \neq t_j$	$d = t_j$	I
0.5	0.5	Pink	$t_i \neq t_j$	$d = (t_i \text{ or } t_j)$	I
0.5	0.4	Black	$t_i \neq t_j$	$d = t_i$	I
0.5	0.4	Black	$t_i = t_j$	$d \neq (t_i \text{ or } t_j)$	III
0.5	0.2	Brown	$t_i \neq t_j$	$d = t_i$	I
0.5	0.2	Brown	$t_i = t_j$	$d \neq (t_i \text{ or } t_j)$	III
0.5	0	Blue-green	$t_i \neq t_j$	$d = t_i$	I
0.5	0	Blue-green	$t_i = t_j$	$d \neq (t_i \text{ or } t_j)$	III

[a] The color of A_i was purple in all conditions.

RESULTS

Table 3 shows the estimates of the cue accuracies made by each subject on the last learning trial. All subjects recognized that the 100 % cue was always accurate, but one (subject 5) failed to recognize that the 0 % cue was always wrong. It is interesting to note that subject 5 used steps of 10 % units to distinguish between cues perceived to be of different accuracies, assigned a different accuracy to all cues, and apparently used the 100 % cue as his reference point for an equal interval scale of the cues. Four of the subjects erred by overestimating the low accurate cues and two erred by underestimating. Note that all of subject 1's errors were underestimations except for the 50 % cues, and all of subject 5's errors were overestimations.

If one compares the accuracy with which subjects learned the cues symmetrically positioned with respect to the 50 % cues (which provided no information about the correct symbol) it appears that the more accurate cues are more accurately perceived

TABLE 3

PERCEIVED ACCURACY OF CUE AFTER 50 LEARNING TRIALS

Color		Blue	Red	Orange	Pink	Purple	Black	Brown	Blue-green
% Accuracy		100	80	60	50	50	40	20	0
	1	1	0.75	0.05	0.75	0.75	0.40	0	0
	2	1	0.85	0.60	0.20	0.30	0.03	0.02	0
Sub-	3	1	0.80	0.60	0.44	0.48	0.28	0.28	0
jects	4	1	0.80	0.60	0.45	0.50	0.40	0.30	0
	5	1	0.90	0.80	0.50	0.70	0.60	0.40	0.30
	6	1	0.75	0.50	0.50	0.50	0.50	0.25	0
Mean		1.0	0.81	0.60	0.47	0.54	0.37	0.21	0.05
S.D.		0	0.05	0.10	0.16	0.15	0.18	0.15	0.11

than the less accurate cues and that accuracies close to 50 % are harder to learn than those near 0 and 100 %. The mean error associated with the 100 % cues was less than that associated with 0 %, the 80 % cue error was less than the 20 % cue error, and the 60 % cue error was less than the 40 % error. The conditions under which each subject learned the cue accuracies differed from subject to subject in the sense that some counted the number of times each cue predicted correctly and some did not. Hence the conditions are not uniform for all subjects. None of the subjects kept a count through the entire set of 50 trials. Subjects 3 and 4 kept counts to trial 25 and then did not change their estimates thereafter. Subject 2 counted to trial 17 but misrecorded a few times. Subject 6 raised his estimate of a cue by 1 point every time that cue displayed the correct symbol and lowered it by 1 point every time the cue was incorrect, but he did not change the estimate if such a change would exceed the 0-100 % limits.

Table 4 shows the experimental conditions experienced by each subject in terms of his perceptions of the cue accuracies, and also shows the ratio of maximizing to nonmaximizing decisions in each perceived condition. All six subjects agree with the 100 % (blue) cue on every trial in which it was used. Four of the five, those who perceived the blue-green cue to be always wrong, followed the inference from this

TABLE 4

NUMBER OF MAXIMIZING[a]: NONMAXIMIZING DECISIONS BY SUBJECTS, COLOR OF CUE, SIMILARITY OF REPORTS AND CUE SITUATIONS[b]

Subject	Table entry	Colored cue shown below and purple cue reports										All[d]
		Agree (Same)			Disagree (Different)							
		Black	Brown	Blue-green	Blue	Red	Orange	Pink	Black	Brown	Blue-green	
1	Ratio	10:0	0:10	0:10	10:0	(7:3)[c]	9:1	(8:2)[c]	10:0	9:1	10:0	58:22
	Sit.[b]	II	IV	IV	I	I	I	I	I	I	I	
2	Ratio	10:0	9:1	10:0	10:0	10:0	10:0	10:0	10:0	10:0	10:0	99:1
	Sit.[b]	III	III	III	I	I	I	I	I	I	I	
3	Ratio	7:3	8:2	10:0	10:0	9:1	7:3	6:4	9:1	7:3	10:0	83:17
	Sit.[b]	III	III	III	I	I	I	I	I	I	I	
4	Ratio	2:8	3:7	10:0	10:0	10:0	8:2	8:2	8:2	9:1	10:0	78:22
	Sit.[b]	IV	IV	IV	I	I	I	I	I	I	I	
5	Ratio	10:0	10:0	(10:0)[c]	10:0	10:0	9:1	9:1	10:0	10:0	10:0	88:2
	Sit.[b]	II	II	II	I	I	I	I	I	I	I	
6	Ratio	8:2	6:4	10:0	10:0	10:0	(4:6)[c]	(8:2)[c]	(3:7)[c]	9:1	10:0	63:7
	Sit.[b]	II	IV	IV	I	I	I	I	I	I	I	
All	Ratio[d]	47:13	36:24	40:10	60:0	49:1	43:7	33:7	47:3	54:6	60:0	469:71

[a] Maximizing decisions are described in Table 1.
[b] Cue situations are described by codes explained in Table 1.
[c] Ratio in parenthesis is number of decisions that agreed with purple report when maximizing decision is to agree with either report.
[d] Entry does not include data shown in parenthesis (see footnote c above).

cue every time the inferences of the blue-green and purple differed. Subject 1 failed to recognize that this cue provided as much information as a 100 % cue, and followed the inference from the purple cue (which he perceived to be 75 % accurate).

TABLE 5

NUMBER OF MAXIMIZING: NONMAXIMIZING DECISIONS BY SUBJECTS AND PERCEIVED SITUATIONS

Perceived situation	Subjects						Total	
	1	2	3	4	5	6	Ratio	Per cent
I	68:2	70:0	58:12	63:7	68:2	69:1	396:24	94
II	10:0				30:0	8:2	48:2	96
III		29:1	25:5				54:6	90
IV	0:20			15:15		16:4	31:39	44
Total	78:22	99:1	83:17	78:22	98:2	93:7	529:71	88

Based on their perceptions of the cues, all subjects showed more maximizing than nonmaximizing behavior. As shown in Table 5, 88 % of the choices maximized expected payoff. A comparison of the different cue situations indicates that most nonoptimizing decisions occurred in Case IV (both displays agreed but disagreement with the reports was required). Fifty-six % of the 70 decisions in this condition were not optimal, whereas in all other conditions, at least 90 % of the decisions were optimal.

Case III also involved a conflict between agreement with the display and maximization. As shown in Table 5, 90 % of the 60 Case III decisions were maximizing.

DISCUSSION

In both the learning of cue accuracies and the decisions based on the cue reports subjects appeared to have greater difficulty dealing with negative predictors (cues with less than 0.5 accuracy). In real life a cue that predicts accurately less than 50 % is generally of little help in choosing the correct course of action except in those very rare instances when only two courses are open. Also, negative predictors present a more difficult cognitive task since one must reverse the display to arrive at the appropriate action.

The results also suggest that conformity pressures exert considerable influence upon decisions, and in fact are sometimes stronger than the maximization pressures (as seen in Case IV). Maximization pressures were stronger than the conformity pressures in Case III where both negative predictors imply the same action, opposite to their reports. In Case IV, although the two cues differ in their implications, the more accurate cue implies conformity with the display, and the less accurate cue implies rejection of both displays. The conformity pressures in Case III are thus weaker than in Case IV.

But even in Case IV, two of the three subjects responded optimally when the less accurate cue was always incorrect, and failed to maximize only when the less accurate cue was correct about 30 %. It thus appears that people are less influenced by conformity the more certain they are about the "correct" response, i.e., the more "potent" the cue.

REFERENCES

BECKER, G. M., STOCKLIN, P. L., VEGH-VILLEGAS, A., AND WICKELGREN, W. A. Working Papers On: Decision making with conflicting information. SP-237. Santa Barbara, General Electric Company, TEMPO, 1963.

BRUNER, J. S., GOODNOW, J. J., AND AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.

COHEN, D., WICKELGREN, W., AND BECKER, G. M. Decision making with incongruent information. SP-234. Santa Barbara, General Electric Company, TEMPO, 1963.

VAN METER, D., AND MIDDLETON, D. Detection and extraction of signals in noise from the point of view of statistical decision theory. Parts I and II. *J. Soc. industr. appl. Math.*, December, 1955, **3**, 192-253; 1956, **4**, 86-119.

RECEIVED: November 11, 1963