THE STRUCTURE OF CUE-CRITERION RELATIONSHIP AND THE DETECTABILITY OF A CUE

HIROSHI AZUMA AND TAKAKO OCHIAI

University of Tokyo

Two types of cue-criterion relationships were differentiated. They were the probabilistic structure in which the cue hits the criterion value at a given percentage and the covarying structure in which the cue values have significant regression upon the criterion value. These two structures were compared with regard to the detectability of cues in a task of detecting and utilizing the relevant cue, the payoff for correct utilization being controlled. Ss were female college students. The probabilistic cue was clearly easier to detect, and the covarying cue was practically unrecognizable when the effect of higher percentage of hit was partialled out. Some of possible limitations, hypotheses and implications were discussed.

Typically, conceptualization studies ask Ss to allot stimulus displays to different values of the criterion variable. There is a set of cue variables, and a stimulus display is defined in terms of the values on these cue variables. A relevant cue is a cue the values of which show a non-chance relationship with the criterion values when stimulus displays are correctly allotted to criterion values.

We may differentiate two types of the relationships between a relevant cue and the criterion. One is the probabilistic structure and the other is covarying structure. The former is such that an optimal decision based on the value of the relevant cue sometimes hits the correct allotment but sometimes doesn't. The percentage of hit is higher than chance, but when it doesn't hit, the amount of miss is no better than chance expectation. When both cue and criterion values are ordinal, the contingency table of these two will show significantly higher frequency only in the cells exactly on the main diagonal, all off diagonal cells being populated uniformly. In the case of the latter structure, the optimal decision based on the relevant cue may or may not lead to the higher percentage of exact hit, but will reduce the expected amount of miss. In a contingency

table of ordered cue and criterion values, the distribution of the cases will be denser near around the main diagonal and will get thinner as it gets further off diagonal. A possible variant of this latter type is a curvilinear rather than a linear regression.

The majority of concept learning studies so far reported have dealt with binary cues and criteria so that there was no need for distinguishing two types given above.

Recently evolving, however, are a number of studies in which both cue and criterion variables are ordinal or intervally scaled. What Carrol has called the functional learning is the learning of the regression function of the criterion value on one or more than one cue variables (Carrol, 1963). Of necessity, the variables are scaled. In now classical study of multiple probability learning by Smedslund (1955), the criterion value was determined by the weighted sum of the cue values. It was a multiple linear regression version of the functional learning. Uhl (1963) called it the interval concept and Azuma and Cronbach (1966 a) proposed to call it the scalar concept. The last authors have also studied another kind of learning to utilize multiple scaled cues (Azuma and Cronbach, 1966 b). In their experiment, there were two relevant cues which were mu-

tually independent. In 75% of the occasions the criterion was determined by the value of one cue and in the rest of occasions it was determined by the other cue value. Although both of the two cues are useful in overall validity, in any one occasion the knowledge of the cue value which is not dominant in that occasion is not only helpful but even disturbing.

Under these circumstances, the distinction of two types of cue-criterion relationships becomes more important. In above mentioned studies on multiple cue utilization by Azuma and Cronbach (1966 a and b), there was a marked tendency that *Ss* notice one of the relevant cues which was more strongly related to the criterion first and long before they noticed the other which held weaker relevance. Think, therefore, of the problem which an *S* faces when he is to notice this first cue before noticing the relevance of another cue. If he is working with a scalar concept learning in which the criterion value is a weighted linear compound of cue values, the relationship of one cue with the criterion is the covarying type, r_{cc} being short of unity. If he is working with a problem in which relevant cues take turns in determining the criterion value, the relationship between one cue and the criterion is the probabilistic type. If there is a difference in the detectability of these two types of cues, therefore, it may very well influence the difficulty of corresponding multiple cue concept learning tasks. Further, if an *S* intuitively imposes either one of the cue-criterion relationship structures on the problem, it will influence the type and the quality of the solution he tends to attain.

Almost all studies so far conducted on the detection of a cue have worked with a cue of the probabilistic type. It is the purpose of the present experiment to investigate the detectability of a covarying type cue comparing it with that of a probabilistic type cue. More specifically, we will construct learning tasks in which the best achievement is attained by consistently

responding according to a single relevant cue. In two of the tasks the cue is relevant only probabilistically and not covaryingly, in one of the tasks the cue is both probabilistically and covaryingly valid and in another task the cue is relevant in the sense of covariation but not probabilistically. The payoff expected of the best strategy, i.e., the reinforcement for the best strategy will be controlled so as to allow the comparison of the effectiveness of two types of cue-criterion relationship structures in making the cue stand out as the relevant cue.

METHOD

Subjects. 57 womens' college sophomores participated as a part of course requirement. They were divided into following four groups: A_1 (14), A_2 (15), B_1 (13) and B_2 (15). The numerals in the parentheses show the numbers of subjects.

Task. Each subject was to work on a booklet which contained 150 frames. Each frame showed four playing card figures; diamond, heart, spade and club. Each figure contained a numeral; 1, 2, 3, 4, or 5. A frame also

19					correct <i>k</i>
20					correct <i>k</i>
21					correct <i>k</i>
22					correct <i>k</i>
23					correct <i>k</i>
24					correct <i>k</i>

FIG. 1. A page of the stimulus booklet.

showed a blank to fill the response in, and a blank to write the announced criterion value in. There were six these frames on a page of the booklet, and a booklet consisted of 25 such pages (Fig. 1). The task of the subjects was to guess the criterion value which the experimenter announced for each frame after the subjects filled their responses to this frame in. The criterion value was also 1, 2, 3, 4 or 5, and the subjects knew through the instruction that figures contained some partially valid cue for inferring the criterion value. Their task was to minimize the cumulative discrepancy between the guess and the criterion value.

The stimulus series. Numbering frames from 1 through 150, the experimenter divided them into six stimulus blocks of 25 frames each and equated some statistical specifications of the cue values for each block. Subjects were not aware of such division. The statistical specifications controlled were as follows:

- 1) The scores of the playing card figures, i.e., the numerals contained, have zero correlation between different kinds of figures.
- 2) For each kind of figure, the scores take on a rectangular distribution.
- 3) No more than three figures in a same frame take on the same score.
- 4) The positions of different kinds of figures were randomized.

The criterion value "k". As it was mentioned before, each frame with four figures was assigned with a criterion value which we call k . For each block of 25 frames, the k values were constructed to satisfy certain statistical specifications which were different for different groups. These specifications were in terms of two statistics, π and $\Sigma \Delta$. The distribution of k over the range of 1-5 was rectangular.

Taking a relevant figure, which is the diamond in our experiment, π is the percentage that the k value of a frame is identical to the number shown in the diamond (hereafter call a diamond score) of the same frame. When π is .52, a subject who always gives the diamond score as the guess for the k of the same frame will hit the k value 13 times out of 25 attempts. The Δ is the absolute difference between diamond score and the k . Thus $\Sigma \Delta$ is the differ-

ence between the diamond score and the k value summed over a block of 25 figures. If $\Sigma \Delta$ is 12 when π is .52, the same hypothetical subject described above will miss the k value by one point 12 times in a block of 25.

If there is no relationship between scores in playing card figures and k , the expected π is .20 and the expected $\Sigma \Delta$ is 40. If a subject follows the lead of either heart, spade or club scores consistently, his π and $\Sigma \Delta$ will be at the level of these chance expectancy. When he follows the diamond scores, however, his π will be higher and/or his $\Sigma \Delta$ will be smaller than the chance expectancy.

Table 1 summarizes π and $\Sigma \Delta$ specified for each group. $\Sigma \Delta$ of Groups A_2 , B_1 and B_2 had some fluctuation from block to block, and Δ of Group B_1 was short of 2.0 by .5. We had to tolerate these irregularities because a better control was impossible due to the limitation in the number of alternative values both k and diamond could take on.

TABLE 1
Specifications of cue-criterion relationships

	π	$\Sigma \Delta$	$\bar{\Delta}$
Group A_1	.76	12	2.0
Group A_2	.50	12 or 13	1.0
Group B_1	.50	18 or 20	1.5
Group B_2	.20	18 or 20	1.0

The last column, $\bar{\Delta}$, indicates the average of the discrepancy between the k value and the diamond score, using the number of the frames in which $\Delta \neq 0$ as the denominator n .

Since figure scores other than the diamond scores were irrelevant, the best the subjects could do was to give consistently the number in the diamond, tolerating some amount of miss. When subjects do this, Groups A_1 and A_2 will be equally successful as the success is judged in terms of $\Sigma \Delta$, and Groups B_1 and B_2 will be less successful. If the success is to be judged in terms of the percentage of hit, however, A_1 is better than A_2 , A_2 is equal to B_1 and B_1 is better than B_2 .

Predictions. We may formulate three theories of cue detection, each leading to a different prediction of experimental results.

1) The detection of a relevant cue is easier when the values of this cue exhibit the minimal deviation from the k value.

2) The detection of a relevant cue is easier when the values of this cue hits the exact k value at a higher percentage.

3) Both the percentage of hit and the amount of deviation contribute jointly in raising the detectability of a relevant cue.

Theory 1) predicts that the detectability of the relevant cue, diamond, will be ordered: $A_1 = A_2 > B_1 = B_2$.

Theory 2) predicts the order:

$$A_1 > A_2 = B_1 > B_2.$$

Theory 3) predicts the order:

$$A_1 > A_2 > B_1 > B_2.$$

The present study has been designed to see which one of these predictions was as a matter of fact valid.

Procedures. The experiment was run in groups. First, the Ss were given a practice series of six trials with frames containing a circle, a square and a triangle in which subjects learned the procedure, and learned how to use the response blank and how to compare his answer with the correct answer which immediately follows his response. Then, the test booklets were distributed. The experimenter gave the instruction which, after giving some idea about how the pages look, ran as follows: "Each one of these four figures contains a number, 1, 2, 3, 4 or 5. Your task is to guess the value assigned to each set of four figures which appears in each frame. We will call this value k . The k is different for different frames, taking on values 1, 2, 3, 4 or 5.

The numbers in the figures should give you some clue for making a better guess. This does not mean that there is a solution by which you can make a perfect score. It is impossible since the clue given is imperfect. We judge you, therefore, in terms of how close your guesses can get. The deviation of your guess from the correct value will be summed up for all 150 trials, and the score will be better if this summed amount of deviation is smaller. There will be a bonus prize for top scorers. Of course you do not know how to guess at the beginning. Gradually through trial and error, you will

learn that a certain way of utilizing cue informations improves the accuracy of your guess. The principle is consistent throughout entire series, that is, a method which is effective at a given part of the booklet should be effective with other parts, too.

Note that the subjects were instructed to try to minimize the deviation.

They worked frame by frame at a controlled rate. The experimenter calls, for example, "frame 7", and subjects will fill in the guess. Then the experimenter shows a cardboard on which the k value is written, and the subjects writes this k value beneath his guess. After everyone is over, the experimenter directs to move to frame 8.

After all 150 frames were over, the subjects were asked to write down the hypotheses they were employing toward the end of the series.

RESULTS

Results are presented in terms of three statistics: The statistic p is the percentage at which a subject gave the same value (numeral) as the diamond score (the relevant cue) as the guessed value of k . The value d is the deviation of a response from the diamond score, and the statistic Σd is d summed over a block of 25 trials. (Note that d is *not* the deviation from k .) The statistic $\bar{d}$ is Σd divided by the number of frames with which the subject *did not* give the diamond score as his guess. The expected value of $\bar{d}$ is 2.00 if the subject's responses were entirely independent of the diamond score for these frames. The statistics $\bar{p}$, $\Sigma \bar{d}$ and $\bar{\bar{d}}$ are the averages of respective statistics over the subjects of a group.

If a subject gives responses completely independent of the diamond score, his $\bar{p}$ will be .20, his Σd will be 40 and his $\bar{d}$ will be 2.00. If a subject always gives a response one point shifted from the diamond score, his p will be 0, his Σd will be 25 and his $\bar{d}$ will be 1.00. If there is a subject who takes the diamond score 10 times in a block but in other occasions gives any value which is not the diamond score, his p will

TABLE 2
Group averages of block by block results

Groups			Learning blocks							
	Conditions		<i>n</i>		1	2	3	4	5	6
A ₁	π	.76	14	$\bar{p}$	39.71	72.29	77.71	80.00	80.85	88.29
	$\bar{d}$	2.00		$\bar{d}$	1.98	1.98	1.64	1.43	1.26	1.52
A ₂	π	.50	15	$\bar{p}$	39.73	46.40	40.53	49.87	50.93	54.67
	$\bar{d}$	1.00		$\bar{d}$	1.95	1.79	1.70	1.66	1.97	1.91
B ₁	π	.50	13	$\bar{p}$	33.42	39.14	46.86	40.86	42.00	46.85
	$\bar{d}$	1.50		$\bar{d}$	1.94	1.79	1.78	1.83	1.86	1.87
B ₂	π	.20	15	$\bar{p}$	28.27	22.40	27.20	23.46	23.73	22.93
	$\bar{d}$	1.00		$\bar{d}$	2.02	1.66	1.71	1.71	1.71	1.70

TABLE 3
Record of individual performances in Blocks 2 and 6

Block 2

Conditions		$\sigma \bar{d}$														
A ₁	$\bar{p}$	92	92	76	88	52	72	84	76	96	76	72	28	44	64	
	$\bar{d}$	2.00	2.50	2.00	2.00	1.75	1.86	1.75	1.67	2.00	2.16	1.72	2.00	2.25	2.00	0.22
A ₂	$\bar{p}$	52	76	80	56	40	28	20	36	40	44	24	72	36	44	48
	$\bar{d}$	1.84	1.00	2.00	2.00	2.33	1.78	1.85	1.87	1.53	1.78	1.79	1.43	1.75	2.07	1.85 0.29
B ₁	$\bar{p}$	28	48	44	36	64	52	44	16	40	44	32	20	48		
	$\bar{d}$	1.56	1.62	2.00	1.82	1.89	2.00	1.86	1.86	1.87	1.57	1.57	2.05	1.62		0.17
B ₂	$\bar{p}$	16	36	28	28	60	20	40	16	12	12	12	16	20	12	8
	$\bar{d}$	1.43	2.00	2.05	1.89	1.70	1.65	1.67	1.62	1.59	1.64	1.64	1.29	1.50	1.82	1.35 0.21

Block 6

Conditions		$\sigma \bar{d}$														
A ₁	$\bar{p}$	100	100	100	96	76	96	88	92	100	88	76	52	92	80	
	$\bar{d}$	0	0	0	1.00	1.83	3.00	2.00	2.50	0	2.67	1.83	1.83	2.50	2.00	0.98
A ₂	$\bar{p}$	64	84	92	84	80	40	20	36	48	24	16	60	48	44	80
	$\bar{d}$	1.78	1.56	2.50	1.75	1.80	1.80	1.65	2.06	2.08	2.21	1.72	2.20	1.54	2.21	1.80 0.27
B ₁	$\bar{p}$	84	80	36	24	96	56	56	28	24	60	12	28	56		
	$\bar{d}$	1.50	2.20	1.75	2.10	2.00	2.18	1.82	2.00	1.84	1.60	1.69	1.78	1.82		0.21
B ₂	$\bar{p}$	28	20	4	44	40	16	20	12	36	36	4	28	4	28	20
	$\bar{d}$	1.39	1.65	1.88	1.50	1.73	1.62	1.70	1.73	1.75	2.12	1.96	1.67	1.52	1.78	1.60 0.18

be 0.40, his Σd will be 30 and his $\bar{d}$ is expected to be 2.00.

Table 2 shows the average block by block results of each group. Table 3 shows individual records of $\bar{p}$ and $\bar{d}$ for blocks 2 and 6, representing beginning and the last stages of learning.

The $\bar{p}$ of Group A₁ goes up rather rapidly as learning proceeds. Toward the end of 150 trials, almost 90% of responses of

this group was exactly the same as the diamond score, indicating that the great majority of the subjects of this group noticed that the diamond score was a highly valid cue in making guess about k . Incidentally, it will be seen that $\bar{p}$ at the last block, 88.29, is much higher than π , .76. This phenomenon of overshooting the probability match expectation was pointed out and discussed elsewhere by one

of the present authors in connection with a different study (Azuma and Cronbach 1966 b). Ten out of 14 subjects also reported verbally of detecting the relevance of the diamond. Also noteworthy is the fact that $\bar{d}$ of this group gets lower than the expected value of 2.00 as learning proceeds. The difference is significant at 5% level for blocks 4, 5 and 6. This means that the subjects tended to give the value close to the diamond score even when they were not giving the guess which is not exactly equal to it. Summing up, most of the subjects of Group A₁ learned the relevant cue and utilized it.

Now we turn to Group A₂. As it will be seen from Table 1, if a subject of this group consistently gives the diamond score as his guess for k , his amount of deviation from the correct k will accumulate to be 12 or 13 at each block. This cumulative amount is about equal to that of a subject of Group A₁ who follows the same principle in giving his guess. Thus, the utility of the response of "giving the diamond score" for Group A₂ subjects is equal to that for Group A₁. Yet, as it is summarized in Table 2, Group A₂ was less effective in developing the tendency to give the diamond score which was the best possible strategy. The $\bar{p}$ of .55 at block 6 is significantly lower than that of .88 of Group A₁. This low $\bar{p}$ should not be interpreted as probability matching to π of .50, although the numerical resemblance may invite our readers to do so. Table 3 shows that the distribution of p of Group A₂ is typically bimodal at Block 6, 13 out of 15 subjects falling either over .80 or under .50. Five subjects who yielded the p of .80 or higher would reasonably be inferred to have picked up the relevance of the diamond score, and three of them did verbally point it out. It is doubtful whether the rest of the subjects of this group had any idea as to the relevant cue, and certainly no one of them verbally expressed it. Therefore, the authors are inclined to attribute the low $\bar{p}$ of group A₂

not to probability matching but to the lower percentage of the subjects who detected the relevant cue.

Another thing to be pointed out with regard to Group A₂ is the fact that $\bar{d}$ consistently stayed higher than $\bar{d}$ which was 1.00. Even though the diamond score was always close to that of k even when they are not equal, the performance of the subjects, in the average, did not reflect this aspect of cue-criterion relationship and there was no one who explicitly pointed it out.

Thus we may conclude that it was much more difficult for Group A₂ than for Group A₁ to detect the relevant cue and to utilize it in spite of the equal payoff expected of the successful detection.

Group B₁ has, as you will recall, worked under a condition which was different from that of A₂ in one aspect. The π was the same as it was with Group A₂, but Σd was larger. Since Σd was what the subjects ought have tried to minimize, the payoff of the strategy of giving the diamond score as the best guess was not quite as great with this group as it was with Group A₂. If the payoff in this sense was the determinant of the detectability of the relevant cue, the detectability should be lower with this group. The results in firm this expectation. The response of Group B₁ was very close to that of A₂ in terms of both $\bar{p}$ and $\bar{d}$. Three out of 13 subjects gave verbal report that they noticed the relevance of the diamond score. Every aspect of the results points to the conclusion that the task of Group B₁ was about equal in psychological difficulty to that of Group A₂.

Group B₂ worked under a condition which was equal to that of B₁ in terms of Σd but had smaller π . The payoff of the strategy of consistently giving the diamond score as the answer is the same as it was with Group B₁.

Yet, Table 2 shows that the performance of Group B₂ was much poorer than that of Group B₁ in terms of $\bar{p}$, and not appre-

ciably different in terms of $\bar{d}$. No one in Group B₂ attained the p higher than 0.50, and no one verbally reported that he detected the relevant cue.

Thus our results show the pattern $A_1 > A_2 = B_1 > B_2$ which is the order predicted by our hypothesis 2) which says "The detection of a relevant cue is easier when the values of this cue hits the exact k value at a higher percentage". Implied by this theory is that it is π and not Σd or $\bar{d}$ which determines the detectability of the relevant cue. Our observation seems to support this generalization. Other theories which considers the average absolute deviation of the cue value from the criterion value as a factor are to be rejected.

DISCUSSION

In conducting any experiment of a new design, there are many decisions which E has to make rather arbitrarily without a definite guarantee that another decision could not have led to another result. So was it with our present experiment.

Followings are some of those arbitrary decisions which we suspect could have influenced our results:

1. The cues were presented in the form of playing card figures. It is highly probable that different individuals had different preliminary sets as to how to utilize these cues depending on the experience they have had with card games.

2. The cue values and k values were always integers. They covered the same range, i.e., 1 through 5. These would have made the development of the hypotheses that k was a weighted or unweighted compound of cue values difficult. When they developed nevertheless, it was hard to test their validity.

3. The instruction directed the S s to take, borrowing the phrases of Bruner et al. (1956), the gambling attitude rather than the problem solving attitude. The experiment was conducted in a classroom set up. A classroom is supposed to be a

place for problem solving and not for gambling. It is very likely that there was a conflict between the attitude directed by the instruction and that directed by the total situation, which could have resulted in a poorer control of the attitude.

All these possibilities can we check by a series of control experiments. It was another arbitrary decision of ours to communicate our findings before taking time in running many more experiments and ask comments and suggestions of the people who are working on related topics.

The major finding which we would like to communicate is that S s were able to detect and utilize a cue which hit the criterion at a percentage well above the chance, but were very inefficient in detecting or utilizing a cue which covaried with the criterion but did not hit the exact value quite as often. It was the percentage of hit which determined the utilization of a cue and the presence or absence of correlation with the criterion did not seem to matter much. Bjorkman (1965) contrasts two types of learning processes; probability learning and functional learning. The former is the learning of the frequencies of the occurrences of events and the latter is the learning of the functional dependence of criterion events on cue values. In the light of the experiments made by Lindahl (1964), he showed a strong tendency for probability learning although a direct comparison of two types was not made. He assumes that differences in frequency are more accessible than ratios and differences between ratios. We may add that differences in frequency are also more accessible than differences in cumulated amount of deviations.

At the present stage, we cannot go very far beyond sheer speculation as to why they are so.

One possibility is that the subjective utility is curvilinearly related to the amount of miss intrinsically. The payoff function which the E set up by his instruction was a linear decay function of the amount of

deviation from the correct k value. The payoff of the strategy of always giving the diamond score was equal for groups A_1 and A_2 , and for groups B_1 and B_2 , only in terms of this E determined payoff function. If the subjective utility is high at zero deviation or exact hit but drops steeply down at the point of minimal significant deviation and flattens out thereafter, the frequency of hit is important and the covariation without increased hit is less important in maximizing such subjective payoff under our present experimental conditions.

May be this is not quite as much a matter of payoff or reward as a matter of search strategy. In testing hypothesis the S does not wait until the end of a series. Instead, he asks for answer from each frame. A 'hit' is a signal to keep the hypothesis at least for the time being, while any amount of 'miss' is a signal to look for possible revision. Although S s were told that there is no perfect cue and a certain amount of miss has to be tolerated, they did not know just how much miss was expectable from the best cue and therefore would not take any miss negligible.

Azuma and Cronbach (1966 a) observed that *S does not solve a single problem* even though E regards the universe as an undifferentiated whole. The same was observable in our present experiment in spite of the fact that in our present study no subdivision of stimuli by S could possibly result in much gain. This tendency is exemplified by post-experiment remarks of a few S s: "In the upper portions of the pages, the heart score was the k and the diamond score dominated in the lower portions". "Whenever there are more than one scores of the same value in one frame, that value is the k value of the frame". Such tendency for subdividing stimuli will make it more difficult to wait and evaluate the cumulative performance toward the end of a series.

Or perhaps there is a simpler explanation like "counting is easier than adding". Even from a crude reflection on the de-

velopmental processes, it will be possible to infer that man is endowed with a better counter and poorer or less accessible accumulator of metric data.

So much for speculations which at least did bring out a number of points to be studied further. Different materials, different methods of feedback and record keeping, and different criteria for success will have to be tried.

Since the detection of a single covarying cue is much more difficult than that of a probabilistic cue, and since the solution of a multiple cue problem generally starts with the identification of the most influential single cue, it is understandable that the learning of a scalar concept is difficult as it impressed Smedslund (1955) and Azuma and Cronbach (1966 a).

SUMMARY

Two types of cue-criterion relationships were defined. They are the probabilistic structure and the covarying structure. The former is defined in terms of the probability that the cue value hits the criterion value, and the latter is defined in terms of how close the cue value stays around the criterion value the percentage of exact hit being partialled out.

The difficulty of four tasks were compared. For all of them, the best performance was possible if S consistently gave the value of one cue as the guess for the criterion value. This relevant cue had to be detected from among four potential cues three of which were irrelevant.

Task A_1 dealt with a highly valid probabilistic cue. Task A_2 was with a cue which was moderately valid probabilistically but, since it also covaried with the criterion, had a low cumulative deviation from the criterion which equaled to that of Task A_1 cue. Task B_1 was with a cue which was mediumly valid in probabilistic sense without additional covarying nature. The cue in Task B_2 was covarying but the percentage of exact hit was at the chance level and the expected cumulative devia-

tion was the same to that of Task B_1 cue.

Fifty seven female college students were divided into four groups and each group worked on one of the tasks.

The order of difficulty was, from the easiest, $A_1 > A_2 = B_1 > B_2$.

It was concluded that people were more sensitive to the probabilistic structure than to the covarying structure of the cue-criterion relationship.

REFERENCES

- AZUMA, H. & Cronbach, L. J. 1966a Cue-response correlations in the attainment of a scalar concept. *Amer. J. Psychol.*, **70**, 38-48.
- AZUMA, H. & Cronbach, L. J. 1966b Concept Attainment with Probabilistic Feedback. In Hammond, K. R. (Ed), *The Psychology of Egon Brunswik*. New York: Holt, Rinhart & Winston.
- BJORKMAN, M. 1965 Studies in predictive behavior. *Scand. J. Psychol.*, **6**, 129-156.
- BRUNER, J. S., Goodnow, J. J., & Austin, G. A. 1956 *A Study of Thinking*. New York: Wiley.
- CARROL, J. D. 1963 Functional learning: the learning of continuous functional mappings relating stimulus and response continua, RB-63-26, Educational Testing Service.
- LINDAHL, M. 1964. The importance of strategy in a complex learning task. *Scand. J. Psychol.*, **5**, 171-180.
- SMEDSLUND, J. 1955 Multiple probability learning: An inquiry into the origins of perception. Oslo: Akademisk Forlag
- SUMMERS, S. A. 1962 The learning of responses to multiple weighted cues. *J. exp. Psychol.*, **64**, 29-34
- UHL, C. N. 1963 Learning of interval concepts: I. Effects of differences in stimulus weights. *J. exp. Psychol.*, **66**, 264-273.

(Received May 15, 1966)