UMEÅ PSYCHOLOGICAL REPORTS

NO. 80 1975

Department of Psychology
University of Umeå

INTEGRATION RULES IN A MULTIPLE-CUE PROBABILITY LEARNING
TASK WITH INTERCORRELATED CUES

BENGT-ÅKE ARMELIUS KERSTIN ARMELIUS

INTEGRATION RULES IN A MULTIPLE-CUE PROBABILITY LEARNING TASK WITH INTERCORRELATED CUES

Armeliuſ, B., and Armeliuſ, K. Integration rules in a multiple-cue probability learning task with intercorrelated cues. Umeå Psychological Reports No. 80, 1975. - The question of how the subjects use the cues in multiple-cue probability learning tasks was studied by having the subjects fill in a questionnaire asking them to describe how they had made their predictions. The questionnaire was given after the subjects had completed their learning of a two-cue suppressor variable task for 100 trials. For 19 of the subjects it was possible to formulate a model on the basis of their verbal report. The models were classified as a) linear models b) configural models or c) estimated weights models. The correlation between the responses generated by the model and the actual responses was computed for each subject. Goodness of fit of the models was found to be quite satisfactory. The results of the learning phase showed that ten subjects reached a performance higher than that expected if they only utilized the information provided by the cue criterion correlations. Performance was highest for subjects using a linear model, while the achievement was low for subjects using an estimated weights model due to the low consistency. The performance of subjects using configural models was relatively poor due to the low validity of the configural models in the present task. When the validity of the models was taken into account, however, the configural models were found to be as easy to follow as the linear models. The conclusions were that it is possible to use the verbal reports given by the subjects to study the strategies employed by the subjects in MCPL tasks, and that it is necessary to do so since very different psychological processes may be expressed in the same mathematical model.

Studies of the effects of cue intercorrelations on inference behavior have focussed mainly on the effects of the cue intercorrelation factor on achievement and its components as described by the so called lens model equation (Hursch, Hammond & Hursch, 1964). The results of these studies show little or no effect of intercorrelations among the cues on performance (e.g., Armelius & Armelius, 1975b; Brehmer, 1974b; Knowles, Hammond, Stewart & Summers, 1972; Naylor & Schenck, 1968).

In an attempt to explain the effects of cue intercorrelations, Boyle (1970) proposed that subjects take only one cue into account at a time when learning the relations between the cues and the criterion. The subjects thus fail to utilize the intercorrelations among the cues, even though they apparently learn about these correlations (Armelius & Armelius, 1975a; Knowles, Hammond, Stewart & Summers, 1972). And as a consequence, their utilization of the cues tends to match the cue-criterion correlations, rather than the cue-criterion beta weights, as would be required for optimal performance. This is an attractive hypothesis, and it is consistent with the results of some of the studies, especially the finding in the experiments by Armelius and Armelius (1975b) and Miller and Sarafino (1970) that the beta-weights for the cues in the regression equations fitted to the subjects' responses are more strongly related to the cue-criterion correlations than to the cue-criterion beta-weights. There is, however, also evidence against the hypothesis. In a transfer study, Brehmer (1971) compared the effects of a change in the cue-criterion correlations with the effects of a change in the cue-criterion beta-weights. The results showed that the subjects changed their utilization of the cues when the cue-criterion beta-weights changed, but not when the cue-criterion correlations were changed. The subjects' utilization of cues did not follow the beta-weights completely but the results are nevertheless evidence against the hypothesis that the subjects' utilization of the cues is determined only by the cue-criterion correlations. Further evidence against the hypothesis is presented in a study by Brehmer (1974b) which also failed to demonstrate a relation between cue-criterion correlations and cue utilization.

These results leave us in a position where we know something about the effects of cue intercorrelations on the subjects' performance in

multiple-cue probability learning tasks, but where we do not know much about the actual utilization of the cues which produces the performance. One reason why these studies have failed to give clear information about how subjects use intercorrelated cues may be that these studies have relied on multiple regression statistics for the description of the subjects' utilization of the cues. Multiple regression is, however, not ideally suited to this purpose since when the cues are intercorrelated, multiple regression does not yield any index of the importance of a cue for the judgment (Darlington, 1968). The present study, therefore, attempts to study cue utilization in a task with intercorrelated cues by means of a different approach, namely through verbal reports from the subjects. This approach has been useful in studies of the subjects' hypotheses about functional relations between cue and criterion in single-cue probability learning (see Brehmer, 1974a; Brehmer, Kylenstierna & Liljergren, 1974), and it is therefore possible that it might provide some insight into how the subjects use the cues also in multiple-cue probability learning tasks.

Method

Subjects. Twenty-four undergraduate students from the University of Umeå participated in the experiment to fulfill a course requirement.

Learning task. The learning task was a two-cue MCPL-task with a suppressor cue. The intercorrelation between the two cues was $r_{ij} = .80$, and the criterion correlation for the valid cue was $r_{e1} = .80$. The suppressor cue was given the criterion correlation $r_{e2} = .30$, which maximizes the suppressor effect, i.e., $R_e^2 = 1.00$. This results in the values of β_{e1} and β_{e2} of 1.56 and $-.94$ respectively.

A suppressor variable task was used because this kind of task maximizes the difference between cue-criterion correlations and cue-criterion beta-weights. Consequently, this kind of task is well suited for studies of what kind of information is used by the subjects.

Procedure. The learning task was presented in booklets. On the face of each page in the booklet the cues were presented as two bars numbered from one

to thirty with the value of each cue represented as the shaded part of the bar. On the other side of the page the criterion value was presented as a number between one and twenty. On each of the 100 training trials the subject observed the two cue values, gave his prediction of the criterion value in his answer sheet and observed the correct criterion value. The subjects were allowed to work at their own pace and were not informed of the structure of the task. They were told to base their predictions on the values of the cues they were shown and it was emphasised that due to the nature of the task they should not expect to be perfectly correct on each trial.

Verbal reports. When learning was completed each subject was given a questionnaire with a number of questions asking them to describe how they had made their predictions. (See appendix.) As shown by Brehmer (1974a) the subjects are able to report the rules they use in a single cue probability learning task. He also found that the subjects' verbal reports give a fairly accurate description of their actual responses.

Results

Description of subjective rules. The subjects' verbal reports were used to define a prediction rule for each subject. For nineteen of the twenty-four subjects, it was possible to extract such a rule. The reports of the remaining five subjects were too incomplete or inconsistent for this purpose. The rules were classified in three categories: 1. Linear models. 2. Configural models and 3. Estimated weight models. Each is described in turn.

1. Linear rule. Some subjects formulated rules that can be written as linear equations similar to regression equations. (See Eq. 1.)
E.g., " $\bar{X}_2$ minus 5 plus 2 for each step of the difference $X_1 - X_2$ ".
Some of the reported rules include complex terms such as the difference between the cue values, whereas other rules were simple averages of the cue values.

$$y' = \sum_{i=1}^n \beta_i Zx_i \quad (1)$$

where y' = simulated response
 Zx_i = standardized cue values
 β_i = subjective beta-weights
 n = number of cues

2. Configural rule. A configural rule was defined as a rule which gives different equations in different parts of the cue matrix. E.g., " X_1 was strong and positive. It determined the value approximately. X_2 varied strongly negatively if it was higher than X_1 and weakly positively if it was lower than X_1 ". The equations in this category were conditional upon the values of the cues, and the weights in different parts of the cue matrix had to be estimated by means of an iterative procedure where different weights were compared.

3. Estimated weight. All subjects who had given an estimated weight in percent for each cue and no complete verbal rule were classified in this category. For tasks where cues are intercorrelated the interpretation of these weights poses a problem, however. Two obvious alternative interpretations are possible. The estimated weights might refer to the beta-weights and to the correlations, and these interpretations were selected for further study. Two simulations were made. In one simulation the estimates were interpreted as correlations and the beta-weights were computed for each subject. The rule was formulated as a regression equation with partial regression weights. In a second simulation the estimates were interpreted directly as beta-weights. Since two subjects gave verbal descriptions which indicated that they were aware of the negative relation between the suppressor cue and the criterion, a third simulation was made for these subjects with the beta-weight for the suppressor cue as negative. The results of all simulations together with achievement, r_a , in the last block of learning are displayed in Table 1.

Table 1. Achievement, r_a , and the correlation between the simulated responses and the actual responses, when the estimated weights (EW) are interpreted as correlations (r) beta-weights (β) and beta-weights with the suppressor-cue negative (β_{neg}).

Correlation between simulated and actual responses				
subject number	r_a	EW=r	EW= β	EW= β_{neg}
1	.18	.06	.59	
10	.59	.76	.96	
12	.67	.66	.54	
13	.69	.72	.62	
16	.77	.83	.76	.39
18	.83	.88	.87	.92
19	.86	.87	.72	
20	.89	.91	.77	
21	.92	.92	.81	

As can be seen from this table the interpretation in terms of correlations produces results closer to the actual responses for six of the subjects. For three subjects the interpretation as beta-weights produces the best correspondence between simulated responses and actual responses. One of the two subjects who noted the negative sign of the suppressor cue has obviously estimated beta-weights, while the other subject has estimated correlations. For further analyses the interpretation that gives the highest correspondence between actual and simulated responses is used for each subject.

Goodness of fit. For each subject, predicted responses were derived for the last 20 trials in learning by means of the model constructed from the subject's verbal report. These predicted values were then plotted against the actual responses of the subject for the corresponding trials, and the relation was subjected to a polynomial regression analysis to investigate whether there were systematic deviations from linearity in the relation between predicted and actual responses.

The polynomial regression analysis showed that the predicted responses account for a large proportion of the variance in most cases. Significant deviations from linearity were found for only four subjects, and for these, the deviations from linearity accounted for only 6% of the variance on the average. Inspection of the plots revealed that the deviations were due to one or two responses for which the model predicted a negative value, but which responses were, of course, positive. Consequently, it seems safe to conclude that in most cases, the models actually account for the subjects' responses, and that their verbal reports thus were accurate. The correlations between the actual responses and the predicted responses are shown together with the performance measures in Table 2.

Expected performance. In order to answer the question whether or not the subjects are able to utilize the intercorrelation between cues to improve their performance beyond the level expected if they only utilize the cue criterion correlations it is necessary to know the expected performance if the subjects are able to utilize r_{ij} and if they are not able to utilize r_{ij} . As shown by Hursch, Hammond and Hursch (1964) the correlation between the subjects' responses and the criterion, r_a , is equal to

$$r_a = \sum_{i=1}^n \beta_{si} r_{ei} \quad (2)$$

where r_{ei} = cue criterion correlation

β_{si} = cue judgment beta-weight

n = number of cues

assuming no nonlinear correlation between the criterion and the judgment.

If the subjects are able to utilize both r_{ij} and r_{ei} and if $\beta_{si} = \beta_{ei}$, then

$$r_a = R_e^2 \quad (3)$$

where R_e^2 = the squared multiple correlation between cues and criterion.

In the present study the expected performance is $r_a = 1.00$ and $\beta_{si} = \beta_{ei}$ if the subjects are able to utilize r_{ij} . If the subjects are not able to utilize r_{ij} but only r_{ei} on the other hand, then $\beta_{si} = r_{ei}$ and

$$r_a = \sum_{i=1}^n r_{ei}^2 \quad (4)$$

The expected performance if the subjects are unable to utilize r_{ij} in the present study is $r_a = .73$ and $\beta_{si} = r_{ei}$.

Individual performance. In table 2 the correlation between the subject's judgments and the correct criterion values, r_a , the squared multiple correlation between cues and judgments, R_s^2 , the cue judgment beta-weights, β_{si} , the type of rule and goodness of fit for each subject in ascending order of r_a for the last 20 learning trials are presented.

Table 2. Achievement, r_a , subject consistency, R_s^2 , cue-judgment β -weights, β_{si} , type of rule and goodness of fit for the models for each subject. The performance measures are taken from the last learning block.

Subject	r_a	R_s^2	β_{x1}	β_{x2}	Type of rule	Goodness of fit
1	.18	.50	-0.03	.73	Estimated weight	.59
2	.28	.38	.10	.53	Linear rule	.60
3	.39	.70	.22	.65	Not specified	-
4	.41	.72	.25	.63	Linear rule	.84
5	.42	.53	.46	.60	Not specified	-
6	.44	.69	.32	.54	Not specified	-
7	.47	.73	.34	.55	Not specified	-
8	.53	.95	.41	.61	Linear rule	.97
9	.56	.71	.60	.28	Configural	.81
10	.59	.93	.55	.46	Estimated weight	.76
11	.65	.81	.75	.18	Not specified	-
12	.67	.44	1.05	-.57	Estimated weight	.66
13	.69	.53	1.06	-0.47	Estimated weight	.73
14	.72	.82	.90	.01	Configural	.85
15	.76	.85	.93	-0.02	Linear rule	.92
16	.77	.73	1.05	-.26	Estimated weight	.83
17	.79	.86	1.03	-0.13	Configural	.83
18	.83	.87	1.16	-.29	Estimated weight	.92
19	.86	.77	1.34	-0.66	Estimated weight	.87
20	.89	.85	1.32	-0.55	Estimated weight	.91
21	.92	.89	1.40	-0.64	Estimated weight	.92
22	.92	.86	1.45	-.77	Linear rule	.93
23	.96	.96	1.70	-1.28	Linear rule	.97
24	.97	.97	1.65	-1.07	Linear rule	.98

As can be seen from Table 2, ten subjects reached a higher level of performance than that expected if they utilized only the cue criterion correlations, i.e., $r_a > .73$. Taken together with the fact that the cue-judgment beta-weights for these subjects deviate from r_{ei} in the direction of β_{ei} , especially for the suppressor cue, these results indicate that the subjects have been able to utilize the beta-weights for the cues, although the cue-utilization is far from optimal. Some of the subjects, however, have been able to find an almost optimal integration rule.

Performance differences between subjects using different rules. Due to the very small number of subjects using each rule any comparisons between the performance of subjects using different rules should be made very carefully. In Table 3 the average performance in the last block of learning for the various rules are presented.

Table 3. Average achievement, r_a , consistency, R_s^2 , and cue-judgment beta-weights, β_{si} , for the different types of rules.

	r_a	R_s^2	β_{s1}	β_{s2}
Not specified	.48	.69	.39	.48
Configural rule	.70	.80	.84	.05
Estimated weight	.77	.72	.99	-.25
Linear rule	.81	.81	.93	-.20

The results in Table 3 show that subjects who were able to report their rules in a form that could be simulated performed better than the subjects who were unable to describe their rules clearly. There is also some evidence in favor of linear models but this is hardly surprising in view of the fact that the task also follows a linear rule. Subjects with configural rules have high consistency but poor weighting of the cues.

The reverse seems to be the case for subjects following an estimated weights model. The subjects following linear rules are relatively successful both with respect to consistency and the weighting of the cues.

Relative performance. Since only a few subjective rules allow the subjects to reach a perfect performance in the present task, i.e., the validities of the rules differ, it is not possible to interpret the performance measures in Table 3 as an expression of how well the subjects have succeeded in utilizing their rule. In order to get an estimate of the performance of each subject in relation to the optimal performance attainable indices given the rule reported by the subject, relative performance were computed for each subject. The optimal performance was determined by correlating the responses generated by the rule with the criterion and computing achievement, r_a , and consistency, R_s for each rule. Relative achievement for each subject was obtained by the ratio of the achievement of the subject to the achievement of the rule. Similarly, relative consistency was obtained by the ratio of the consistency of the subject to the consistency of the rule.

An estimate of the extent to which the beta-weights of the subjects correspond to the beta-weights of the rule, i.e., relative matching of regression weights can be obtained by the ratio of relative achievement to relative consistency. The average relative performance for the different rules is shown in Table 4.

Table 4. Average relative achievement, consistency and matching of regression weights for the different types of rules.

	Relative r_a	Relative R_s	Relative matching
Configural rule	.89	.95	.94
Estimated weight	.77	.85	.91
Linear rule	.84	.89	.94

As can be seen from Table 4 the relative performance for the subjects using configural rules is as good as the performance of the subjects using linear rules. This means that the configural rules are as easy to follow as the linear rules although they are not as valid in the present task. The lowest relative performance is reached by the estimated weights. The main reason for this seems to be the low relative consistency.

Discussion

Achievement in the present study was higher than that expected from earlier studies with suppressor variables (Armeliu & Armeliu, 1975b; Miller & Sarafino, 1970). Although the average r_a was far from optimal it was higher than the level expected if the subjects were able to utilize only the cue-criterion correlations and not the intercorrelation when making their predictions. Further support for the conclusion that at least some subjects were able to utilize the information provided by the suppressor cue is found in the negative beta-weights between the subjects' responses and the suppressor cue. One possible explanation for this unexpected result is that the cue-criterion correlations are of importance for the utilization of r_{ij} . In the present study the cue-criterion correlations of both cues were higher than in previous studies with suppressor variables.

The goodness of fit for the verbal reports was high and it may be concluded that when subjects are able to give verbal reports these can be used to describe how they make their predictions. The analysis of the verbal reports shows that those subjects who are able to describe what rules they used when making their predictions also perform better than those subjects who could not describe how they made their predictions. It seems natural that it is easier to revise strategies and test hypothesis in a relatively difficult task if the subjects verbalize the hypothesis they are working with. Whether the verbalization is a requirement for or a result of learning of the task is an interesting question that cannot be answered in the present study.

In general, linear equations may be used to describe the rules used by the subjects. It is important to notice the difference in performance between the rules generated from the subjects estimation of weights and the rules called linear rules. When subjects estimate weights their consistency is much lower than when they are able to formulate a rule that is essentially deterministic. This might be due to a more intuitive set among the subjects working with weights and a more analytical set among subjects working with deterministic rules.

The subjects using configural rules do not reach as high a performance as the subjects using other rules in the present task. The analysis of relative performance shows, however, that the main reason for the low performance is due to the inadequacy of the configural rules in the present task rather than to the difficulty of following such rules. The subjects who have used configural rules reached the same relative performance as the subjects using linear rules. This may be due to the fact that configural rules in the present study may be conceptualized as a set of deterministic rules.

As shown in the present study it is important to distinguish between reported rules and the mathematical expressions for these reported rules. This means, a) that some of the reported rules are much more complex than the mathematical expression shows, e.g., subject number 24 said: " X_2 minus 5, plus 2 for each step of the difference $X_1 - X_2$ " which in regression terms is $2X_1 - X_2 - 5$, and b) it is not enough to analyze the regression equations to describe what rules the subjects use (Naylor & Wherry, 1965; Wherry & Naylor, 1966) and c) two different verbal rules may give essentially the same mathematical equation. In the present study subjects 23 and 24 gave two very different verbal descriptions of their rules, and yet the mathematical equations were quite similar. Subject number 23 reported the rule: "2 points for each step of X_1 minus 1 point for each step of X_2 ", while subject 24 reported the rule described above. As is easily realized these two rules differ only with respect to the intercept in mathematical terms, although the rule reported by subject 24 includes a difference between the two cue values. It is quite evident that there are different psychological processes behind the same paramorphic representation in terms of mathematical equations (Hoffman, 1960).

In summary, the results of the present experiment suggest that it may be quite fruitful to ask the subjects how they make their predictions in MCPL tasks. This approach may perhaps throw some light on the intriguing question why some MCPL tasks, e.g., suppressor variable tasks, are very difficult to learn for the subjects.

This study was supported by a grant from the Swedish Council for Social Sciences Research. The authors are indebted to Dr. Berndt Brehmer for valuable comments on this paper.

Appendix

Below you will find a number of questions concerning your behavior during the experiment and your opinion about how the task was constructed. By means of these questions and your answers we want to try to understand the results of the experiment.

First we want to know what method (s) you have used. Fill in only those alternatives that suit you below.

- a) I gave % weight to bar A (the left) throughout
 % weight to bar B (the right) throughout

- b) I tried different weights during the experiment. In the beginning
I gave % weight to bar A and % weight to bar B,
but later I gave % weight to A and % weight to B.

- c) I gave % weight to bar and then I modified my answer
depending on

- d) My answer was sometimes/completely determined by the combination of
the two bars in the following way:

- e) None of the methods above suits my behavior, but I answered in the
following way:

References

- Armeliu s, K., & Armeliu s B. Note on detection of cue intercorrelation in multiple-cue probability learning. Scandinavian Journal of Psychology, 1975a (in press).
- Armeliu s, B., & Armeliu s, K. Utilization of redundancy in multiple-cue judgments: Data from a suppressor variable task. American Journal of Psychology, 1975b (in press).
- Boyle, P. A comment on parameter matching in MCPL-tasks, 1970. Unpublished paper.
- Brehmer, B. Cue utilization in multiple-cue probability learning tasks with intercorrelated cues. Unaa Psychological Reports, No. 45, 1971.
- Brehmer, B. Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. Organizational Behavior and Human Performance, 1974a, 11, 1-27.
- Brehmer, B. The effect of cue intercorrelation on interpersonal learning of probabilistic inference tasks. Organizational Behavior and Human Performance, 1974b, 12, 397-412.
- Brehmer, B., Kuylenstierna, J., & Liljergren, J-E. Effects of function form and cue validity on the subjects' hypotheses in probabilistic inference tasks. Organizational Behavior and Human Performance, 1974, 11, 338-354.
- Darlington. R.B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.
- Hoffman, P.J. The paramorphic representation on clinical judgment. Psychological Bulletin, 1960, 47, 116-131.
- Hursch, C., Hammond, K.R., & Hursch, J.L. Some methodological consideration in multiple-cue probability studies. Psychological Review, 1964, 71, 42-60.

- Knowles, B.A., Hammond, K.R., Stewart, T.R., & Summers, D.A. Positive and negative redundancy in multiple cue probability tasks. Journal of Experimental Psychology, 1972, 90, 158-159.
- Knowles, B.A., Hammond, K.R., Stewart, T.R., & Summers, D.A. Detection of redundancy in multiple cue probability tasks. Journal of Experimental Psychology, 1972, 93, 425-427.
- Miller, M.J., & Sarafino, E. The effects of intercorrelated cues on multiple probability learning. Program on Cognitive Processes Report No, 128. Institute of Behavioral Science, University of Colorado, 1970.
- Naylor, J., & Schenck, E.A. The influence of cue redundancy upon the human inference process for tasks of varying degrees of predictability. Organizational Behavior and Human Performance, 1968, 3, 147-161.
- Naylor, J., & Wherry, R. The use of simulated stimuli and the "jan" technique to capture and cluster the policies of raters. Educational and Psychological Measurement, 1965, 25, 969-986.
- Wherry, R., & Naylor, J. Comparison of two approaches - jan and prof- for capturing rater strategies. Educational and Psychological Measurement, 1966, 26, 267-286.