HUMAN FACTORS, 1967, 9 (5), 439–446

Confusion Matrix Analysis for Form Perception

RAYMOND D. ENGSTRAND *and* GEORGE MOELLER, *Submarine Medical Research Laboratory, Naval Submarine Medical Center, Naval Submarine Base New London, Groton, Connecticut*

The Constant-Ratio Rule (CRR), an empirical technique for analysis of confusion matrices, was developed for use in predicting intelligibility of speech syllables. This study investigated the validity of the rule when applied to the data from experiments on visual form perception. English letters and simple geometric figures were tachistoscopically presented in the center of a viewing field. Response proportions for subsets of this master set of stimuli were predicted by CRR. Results indicated that the rule (1) accurately predicted numeric response proportions for subsets of stimuli when experimental conditions were similar and (2) predicted ordinally accurate data when experimental conditions varied within the limit which might be encountered in "operational situations." These results, as well as arithmetic factors which can result in errors in prediction, are discussed.

Ideally the stimuli chosen for a display code should form the most discriminable set of all those which might have been constructed. If a parent population has been selected, e.g., phonemes, and the number of signals required has been fixed, how is the ideal set to be chosen? The traditional approaches to this problem are based on either, or both, of the following assumptions: (1) the elements of the population may be ordered with reference to a single psychological dimension, and (2) the distribution of the psychological process(es) activated by each of the stimuli takes some fixed form. For most practicable coding systems the first assumption is false, e.g., phonemes; or it cannot be realized economically, e.g., a set of stimuli varying only in hue would be very expensive to produce. The second assumption cannot be proved or disproved in most cases. The sophisticated multi-dimensional scaling methods under intensive development ultimately may be useful in establishing stimulus dimensionality (Shepard, 1963).

In 1957, F. R. Clarke introduced the Constant Ratio Rule (CRR), an empirical technique for predicting the intelligibility of subsets of speech syllables from a master set of syllables transmitted in a background of noise. The rule was simply the formula for predicting probability values of a truncated distribution. It required only that subject's discriminations be independent of each other. As such, CRR offered the possibility of allowing an experimenter to test a large set of stimuli under conditions similar to operational requirements, predict the discriminability of subsets of stimuli within the master set, and then to select some particular subset of these stimuli for use on specific displays. Bowen, Andreassi, Truax, and Orlansky (1960) used the rule in this manner to select symbols for use on radar displays.

Predictions by the rule are obtained in the following way. The results of an empirical test of stimulus discriminability for some specified master set of stimuli are cast in a confusion matrix.[1] Within the matrix obtained frequencies of response are taken as estimates of the probability of a response to a stimulus. Then, the investigator selects a subset of these probabilities to derive a particular sub-matrix. CRR asserts ". . . that the ratio between any two entries in a row of a submatrix is equal to the ratio between the corresponding two

[1] A confusion matrix is a table listing stimuli along the vertical axis, responses on the horizontal axis, and obtained frequencies of correct and incorrect responses in appropriate cells of the matrix.

entries in the master matrix" (Clarke, 1957, p. 715). It is relatively simple to derive any of the possible confusion matrices contained in a master matrix by application of the following formula to each of the selected subsets of cells

$$p(a_k/Ai) = \frac{P(a_k/Ai)}{\sum_{k=1}^{s} P(a_k/Ai)}$$

where a_k is a response to a stimulus A_i, $p(a_k/A_i)$ is an entry in a submatrix, $P(a_k/A_i)$ is an entry in the master matrix, and s the number of elements in the subset. Application of this formula exactly as described by Clarke, presumes a square matrix. Accordingly subjects would respond on every trial, know the set of stimuli from which each stimulus is drawn, and limit their responses to responses associated with items in this set. There are no standard criteria for selection of a particular submatrix from among all those of a given size which might be derived from a master matrix. Clearly, it is the practical situation which will govern the criteria of the "best" matrix that must be established, and these criteria must take into account the many facets of the intended application in addition to the numerical values of submatrix entries predicted by CRR. For example, it is the characteristics of the equipment, the operator, and the task objectives which govern whether an investigator will be interested in selecting stimuli with the highest probability of identification, or stimuli with a high probability of identification and other characteristics. For more detailed descriptions of the rule see Clarke (1957) or Hodge and Pollack (1962).

The few formal tests of CRR conducted to date have tended to substantiate its predictive usefulness. Clarke and Anderson (1957), Clarke (1957, 1959), Pollack and Decker (1960) and Hodge and Pollack (1962) successfully predicted performance on auditory tasks. Hodge, Piercy, and Crawford (1961) used the rule to predict performance in a weight lifting situation, and Hodge, Crawford, and Piercy (1962) were able to adequately predict performance on a visual task of discriminating differences in the areas of circles. In these investigations confusion matrices derived by the rule compared favorably with empirically obtained matrices for the same subset of stimuli.

In the original development of the rule, Clarke predicted discriminability of sets and subsets of stimuli presented under the same experimental conditions. The purpose of the present study was to investigate the validity of CRR for a form perception task, a type of discrimination to which the rule has been applied without explicit validation, when the conditions of stimulus presentation are varied.

METHOD

Two studies are reported: the first entails a comparison of results from two separate experiments; the second is based on a single experiment. Study I consisted of two experiments designed to investigate the "relative legibility" of upper case English letters and 10 simple geometric figures. Study II consisted of a single experiment designed to assess the ability of CRR to predict the discriminability of a particular subset of the visual forms used in Study I.

Experiment I of Study I was conducted using the 26 letters of the alphabet and 10 geometric figures as visual stimuli presented under two levels of task difficulty (brightness contrasts) and two modes of stimulation (light figures on dark surround and dark figures on light surround). In this experiment the master stimulus set consisted of 36 elements. Experiment II was conducted using 10 letters of the alphabet and the same 10 geometric figures (Figure 1) presented under different levels of task difficulty and the same modes of stimulation. In experiment II the master set of stimuli consisted of 20 elements.

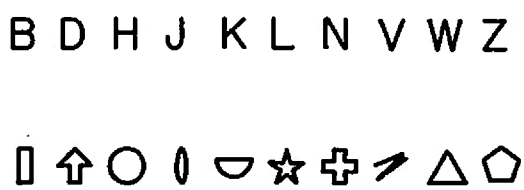

Figure 1. The 10 letters and 10 geometric figures used in experiment II.

The purpose of experiment I was to obtain empirical data from which a subset of 10 letters could be chosen which "matched" the 10 geo-

metric figures in discriminability. Derivation of confusion matrices by CRR was one aspect of the analyses leading to the final selection of the subset of letters shown in Figure 1. In conjunction with empirical examination of the 36 × 36 element matrix, several 20 × 20 element submatrices were considered. CRR was applied to each of the subsets and the numerical data obtained was used in the final comparison and selection of 10 letters.

Experiment II provides empirical evidence on the quality of the match of the subset of letters chosen and the geometric figures. In addition to the check on the "goodness" of the match, empirical and derived confusion matrices obtained in experiment II were compared with corresponding matrices obtained from the data in experiment I. These comparisons provide some evidence on the validity of CRR for prediction across subjects and conditions of stimulus presentation. For this paper, analysis of these comparisons is the purpose of Study I.

Study II consisted of a single experiment designed to assess the validity of a particular subset of response proportions predicted by CRR. On two of four consecutive days subjects were presented the 20 element set of stimuli used in experiment II. On the other two days they were presented the 10 element subset of geometric figures. In Study II all stimuli were presented as dark figures on a light surround and at one level of brightness contrast. Confusion matrices for the 10 element subset of geometric figures were then derived from the 20 stimulus data and compared with the empirical data obtained for the 10 element set.

In all experiments single stimulus characters were presented tachistoscopically in the center of a viewing field, and the subject's task was to recognize items presented from an available listing of stimuli.

Subjects

U. S. Naval enlisted candidates for basic submarine school served as subjects. In Study I there were 12 subjects in experiment I and 28 subjects in experiment II. In Study II there were 10 subjects. All subjects met the physical and mental requirements for admission to submarine school including normal, or corrected to normal, visual acuity.

STUDY I

Apparatus and materials. In experiment I the visual stimuli consisted of the 26 upper case letters of the English alphabet and 10 simple geometric figures. Figure 1 shows the visual stimuli used in experiment II. These are the same 10 geometric figures and 10 of the upper case letters used in experiment I. All stimuli were produced with a stroke width to height ratio of 1:7. Letters were drawn in the "Leroy lettering guide" type style. The figures were drawn with a similar custom-made template. The line drawings were reproduced photographically as high contrast sets of positive and negative 2″ × 2″ slides. The slides used in experiment II were subsets of the slides used in experiment I.

Stimuli were presented in the center of a visual field by means of a two field projection tachistoscope. Exposure time was 100 msec. All stimuli were projected on a background brightness of 7 Ft-L. and produced a retinal image subtending a vertical visual angle of 1° 30′ for subjects seated 7 feet from the visual field. Two levels of brightness contrast conditions were used in each experiment. In experiment I the high brightness contrast condition was a 3.6% change in the brightness of the figure on the background and the low brightness contrast condition was a 2.2% change. In experiment II the high brightness contrast was 3.2% and the low brightness contrast 2.5%. Brightness contrast was controlled by placing neutral density filters in front of the projection lens.

Procedure. The subjects were tested in groups of four, each group serving approximately one hour a day on each of four consecutive days. Each subject served in one stimulus mode condition and both brightness contrast conditions. Each day the subject viewed the complete set of stimuli twice at both high and low brightness contrast for a total of eight responses/brightness contrast/stimulus. Order of the brightness contrast conditions was counterbalanced within and over days. The order of stimuli was random and different on each of the four days, but the same for all subjects.

Stimuli were presented at approximately 10 sec. intervals during a cycle through the entire set of 36 (20) stimuli. Between cycles there was a 1 min. rest period. The subjects were instructed to fixate

a dot in the center of the field on a "ready" signal given 1–3 sec. prior to the presentation of each stimulus.

A multiple choice task was employed, and subjects were instructed to guess when uncertain. Some subjects did not respond to all stimuli; these failures to respond were categorized as "blank" responses.

The occasional failures of subjects to respond creates a problem in computation of confusion matrices insofar as the derivation of the rule assumes a square matrix. Since the number of failures to respond was very small for any single subject, the primary computations were completed by simply omitting such instances from the initial tabulation, i.e., reducing the total possible number of responses per stimulus. Two checks for bias introduced by the modification of the computational procedure were employed. First each set of computations was repeated with all failures to respond assigned to a dummy response category resulting in matrices of 36 rows by 37 columns in experiment I and 20 × 21 in experiment II. Second, in order to obtain a square matrix a small group of subjects was tested with the original set of stimuli augmented by trials in which a blank slide was presented as a stimulus.

Results. Confusion matrices were constructed from the record of each subject's responses to each of the stimuli presented. Separate analyses of variance were performed on the frequency of correct response data for each experiment to assess the effects of the manipulated variables (Engstrand and Moeller, 1962). CRR was applied to the empirical data of experiment I to predict response probabilities for the 20 element matrix of experiment II. The empirical data obtained in experiment II provided a partial test of the accuracy of these predictions.

Figure 2 shows the scatter diagram of the correct responses obtained from the 20 × 20 empirical matrix plotted on the derived data from the 36 element stimulus set of experiment I. The equation for the straight line fitted to the data is $y = .79x - .02$. The correlation between predicted and obtained proportions of correct response is reliably different from zero ($r = .81$). The slope of the equation does not differ significantly from unity. The values shown in Figure 2 are from summary data in which both conditions of brightness and

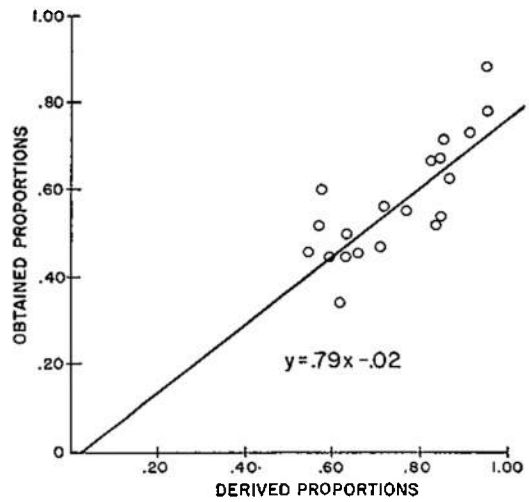

Figure 2. Proportions of correct responses for 20 x 20 element confusion matrix.

both conditions of mode have been combined. Similar comparisons were made for each pair of conditions of stimulus presentation, e.g., stimuli in the same mode but at different levels of brightness contrast, and for each combination of conditions both within and across experiments I and II. Thus, there were 36 comparisons in each experiment and 81 comparisons across experiments for a total of 153 comparisons.[2] The analysis of the combined summary data is presented as representative of all the comparisons.

The results obtained by plotting the obtained correct response values of the 20 × 21 matrices (blanks included) on values derived from the 36 element matrix are essentially identical to those described above. In these two experiments the occurrence of the "blank" response and the consequent departure from a strictly closed set of stimuli and responses did not unduly affect the predictive power of the rule. The exploratory experiment in which "blank" stimuli were included as stimuli also supported that conclusion.

CRR was also applied to the data of experiments I and II to predict 10 × 10 matrices com-

[2] Tables of these intercorrelations have been deposited with the American Documentation Institute. Order Document No. ADI9671, remitting $1.75 for 35-microfilm or $2.50 for 6 by 8 in. photocopies.

mon to both sets of data. There was no empirical data for the 10 × 10 matrices in these experiments. Table 1 shows predicted correct response proportions in the summary matrices for the 10 letter and the 10 geometric figure subset common to both experiments. The left half of the table shows the 10 letter subset, the right half the geometric figure subset. Columns 2 and 5 show the proportions derived from the 36 element data, columns 3 and 6 the proportions derived from the 20 element data. It can be seen that while there are numeric differences in the values of the proportions predicted, the rank orders of the derived proportions from the two experiments are in substantial agreement (r, letters = + .617, $p < .05$; r, figures = + .918, $p < .01$). As for the 20 × 20 data, the various within and across experiment correlations for both alphabetic and geometric subsets were computed, a total of 306 correlations.[2] As a rule the correlations for the 10 × 10 alphabetic matrices across experiments were appreciably lower than for the 10 × 10 geometric matrices and for the 20 × 20 matrices. Within experiments the correlations for the 10 × 10 alphabetic matrices were as high as those for any other comparisons.

Because the preceding analysis has been concerned with the proportions of correct responses, the bulk of the values from each confusion matrix, the elements lying in the triangular matrices above and below the negative diagonal, have been omitted purposefully from the analysis presented here. The mass of those points lie near the origin of the scatter diagram.

Thus the data of Study I demonstrates that: (1) response proportions for a 20 element matrix predicted from the data of a 36 element matrix compare favorably with the empirically obtained proportions of responses to the same set of stimuli presented under different brightness conditions and to different subjects; (2) response probabilities for 10 × 10 submatrices predicted from data collected under different sets of experimental conditions produced similar rank orderings of stimulus discriminability. Discrepancies in the predicted response proportions began to appear when CRR was applied to the data for alphabetic figures obtained under the different conditions of experiment I and II. The reasons for these discrepancies could not be determined from the data of this study.

STUDY II

Apparatus and materials. The visual stimul consisted of the 10 letters and the 10 figures shown in Figure 1. All stimuli were drawn in black india ink on a white background (positive mode) with a stroke width to height ratio of 1:7. Letters were drawn in the "Leroy lettering guide" type style, and figures were drawn with a similar type template. The drawn stimulus characters were ½″ high and varied in width from ¼″ to ⅝″, depending on the form of the character. Stimulus materials were presented by means of a Scientific Prototype Model G Tachistoscope. The exposure interval was 40 msecs. and the brightness of the subject's viewing field was set to .035 Ft-L with a Luckish Taylor Brightness meter. The subject's binocular viewing field was 8° 21′ wide and 5° 34′ high. Stimuli were presented in the center of the visual field. The characters subtended a visual angle of 39′ in height and varied in width from 20′ to 49′.

Procedure. Subjects were tested individually on four consecutive days. On two of the days the subjects viewed only the geometric set of figures (Condition 10). During the other two test sessions the subjects viewed both letters and geometric figures (Condition 20). The order of stimulus conditions over days was either 10, 20, 20, 10 or 20, 10, 10, 20.

A stimulus trial consisted of the following sequence of events. Between trials the subject viewed a recognition field consisting of the 20 possible stimuli. The beginning of a trial was signified by a buzzer sounding for 1 second. Coincident with the buzzer a fixation point in the center of the viewing field was substituted for the recognition field and remained in view for 2 seconds. The visual form then appeared for 40 msec. followed by the reappearance of the fixation point (1 second) and the recognition field in sequence.

Subjects responded verbally at any time following the presentation of the stimulus using words of the international phonetic alphabet for the letters and descriptive names for the geometric figures (circle, star, zig-zag, etc.). Subjects were trained to use the phonetic alphabet and names prior to the start of the test sessions.

All letters and figures appeared an equal number of times during each test session. On the days

when the set of geometric figures was presented, each character appeared 50 times. On the days when letters and figures were both presented, each character appeared 25 times. Thus each subject viewed 500 stimuli on each of four days for a total of 2000 test observations. A single listing of 500 stimuli was prepared for each condition (10, 20). Order of stimuli within a list was random with the restriction that a character could not appear more than four times consecutively. All subjects viewed the same list on both days assigned to a given condition. Data were collected on all four days, but only the records of the last two days were used in analysis of results.

Results. Confusion matrices were constructed from the records of each of the subjects for both the 20 element set of letters and figures and the 10 element subset of geometric figures. Figure 3 shows the mean percentage of responses obtained plotted on the mean percentage of predicted responses for the geometric figure subset. The linear equation fitted to this data is $y = 1.01x$. The Pearson r for these data was .99; r for correct responses only was .83 ($p < .01$).

Figure 4 shows the empirical data obtained in experiment II of Study I plotted on the empirical data obtained in Study II for the 20 element master matrix common to both. This figure demonstrates the reliability of the legibility of these stimulus materials over the conditions employed in these two studies. The line fitted to these data is $y = .76x + .08$. The slope is not significantly different from 1. The Product moment correlation between the two sets of data is $r = .64$ ($p < .01$).

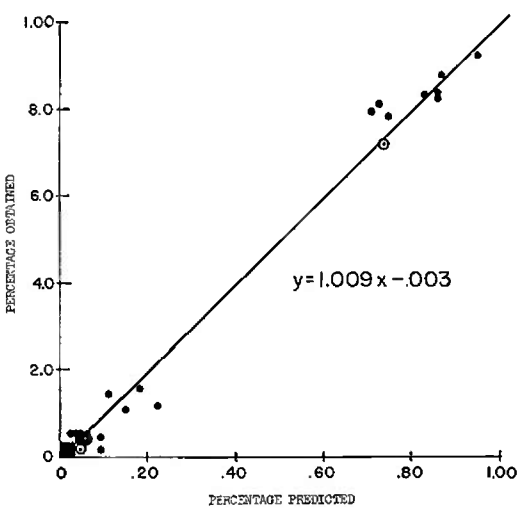

Figure 3. Mean percentage of responses obtained in study II plotted on mean percentage of responses predicted by CRR.

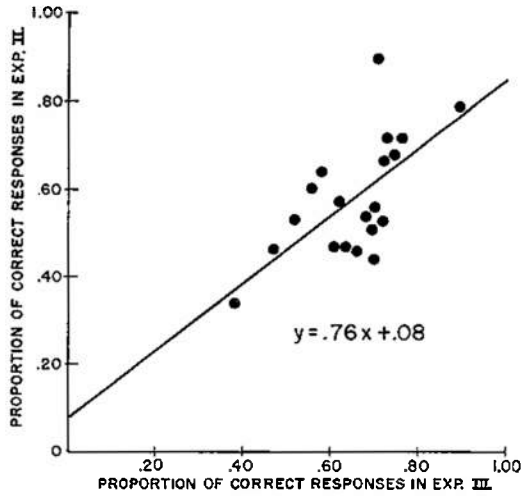

Figure 4. Empirical proportions of correct responses obtained in experiments II and III.

DISCUSSION

CRR was developed by Clarke to predict speech patterns transmitted in a controlled environment. It has been tested in other psychophysical situations, both single and multidimensional, and it seems to have withstood these tests at least as well as other techniques, and better than some. The rule as developed presupposes that the experimental situation under which the master matrix of stimuli is presented (and from which the submatrices of response proportions are to be predicted) will be the same as the experimental situation under which the submatrices are to be presented.

We were concerned with the ability of the rule to predict the discriminability of a set of stimuli which were to be presented under conditions different from those of the original data collection. In all too many cases numeric data obtained under carefully controlled laboratory conditions and analyzed by traditional procedures has had marginal value for generalization to situations other

than those used in the initial data collection. Since CRR is an empirical technique independent of psychological theory, it was hoped that the rule would prove applicable to prediction of response proportions without regard to wide variations in experimental conditions. Under such circumstances the proportions derived could not be expected to be absolutely accurate, because of differences in transmission situations. However analysis of predicted proportions, as ordinal data, might be of value as an aid to the selection of stimuli in practical situations.

Over the range of conditions tested in Study I, the empirical data obtained for the 20 stimuli used in experiment II compared very favorably with the predictions by CRR from the data of experiment I when the latter are taken as indices of relative rather than absolute discriminability of the visual forms. For both experiments I and II the intra-experiment correlations between predicted correct responses were generally positive and significant. Across conditions and experiments there was general agreement in the predictions by the rule for the geometric figures. But across experiments there was not always agreement among the predicted responses for alphabetic forms. It should be noted that none of the discrepancies between the predictions of alphabetic and geometric forms can be accounted for in terms of level of discriminability (letters were slightly better recognized than figures), or differences in conditions or subjects (accurate prediction of geometric forms held across conditions and subjects).

It should be emphasized that the predictions by the rule to interexperimental data were accurate so far as relative discriminability of forms was concerned. Figure 2 shows that the values derived from experiment I for the 20 element set of forms are generally greater than the values obtained empirically in experiment II. Examination of Table 1 also shows that values derived from experiment I for the 10 element subset are larger than the corresponding values derived from the data of experiment II. This apparent tendency of the rule to overpredict, as size of subset (relative to size of master matrix) decreases, has been observed by other investigators.

In experiment I and II the tendency toward overprediction is clearly confounded with differences in task difficulty. In the second of these experiments, the task levels were chosen to be intermediate between those for the first. Study II (experiment III) was conducted in part to obtain further evidence on the overprediction question. All experimental conditions were held constant except for the size of the stimulus set. Data were obtained in separate sessions for the 20 element set used in Study II and the 10 element subset of geometric figures. Figure 3 shows the mean proportion of responses obtained for each of the 10 geometric figures plotted on the proportion of responses predicted by CRR. In this case, CRR tended to predict exact proportions of responses.

TABLE 1

Derived proportions of correct responses for the 10 letters and 10 figures common to both experiments

	LETTERS			*FIGURES*	
	36 Element Data	*20 Element Data*		*36 Element Data*	*20 Element Data*
Delta	.97	.81	Circle	.98	.94
Victor	.95	.76	Triangle	.96	.82
Kilo	.92	.71	Semi-circle	.94	.74
November	.91	.63	Pentagon	.88	.76
Whiskey	.89	.77	Rectangle	.83	.64
Lima	.86	.70	Ellipse	.76	.72
Zulu	.86	.67	Arrow	.68	.51
Bravo	.78	.52	Cross	.67	.52
Juliett	.69	.72	Star	.66	.52
Hotel	.64	.56	Zig-zag	.64	.59

Other authors have commented on deviations in the rule's ability to predict the linear function $f(x) = x$. All arguments have, of course, concluded that an empirical rule which does not provide for changes in parameters with changes in conditions cannot possibly make numerically accurate predictions. Hodge and his colleagues attributed weaknesses in prediction to inter-stimulus spacing (Hodge, 1962), level of task difficulty (Hodge, Piercy, and Crawford, 1961), response-response confusions (Hodge, Crawford, and Piercy, 1961), and practice effects (Hodge, Piercy, and Crawford, 1961).

Clarke (1957) noted that the rule tended to overpredict cells with "large" initial probability estimates and underestimate cells with "low" initial probability estimates. From basic probability theory it can be shown that these tendencies may result from built in bias dependent on (1) the number of observations in each row of a submatrix, (2) the number of stimuli (rows) in both the master and submatrix, and (3) the shape of the

true distribution of responses within a row. In the formula for CRR the numerator is the frequency of observed responses to a given stimulus in the master matrix, and the denominator is the summed frequencies of correct and incorrect responses to the elements of the subset of stimuli. For all derived matrices containing a particular stimulus, the numerator for the proportion of correct responses will be some constant large value. The numerator for the proportions of each incorrect response will generally be some constant small value. The value of the denominator will vary with selection of elements in the submatrix with the usual case being a large entry of correct number of responses and several entries of relatively few or zero observations. On the usual assumption that random error in measurement will be reflected in each observed correct and incorrect response, the greater the discrepancy between the number of cells in the master matrix and the number of cells in the submatrix, the greater should be the error in the estimate of the "true" value of the denominator. If the number of observations in each row is small, then low probability events are unlikely to have occurred. Consequently the high frequency events will provide the major contribution to the denominator. In the case of estimating the probability of a correct response this factor will result in overprediction. In the case of estimating the probability of an incorrect response this factor will result in underprediction.

One additional comment about the accuracy of prediction from the rule ought to be made. The data for the master matrices of experiment II and experiment III were compared to assess the reliability of the discriminations studied. While the product moment correlation between the two master matrices is significant ($r = .64$, $p < .01$), the reliability of the direct discriminations is low for measures to be used as predictors. In view of the relatively low reliability the correlation between obtained and predicted matrices in Studies I and II is quite respectable. Ordinarily we would expect that the test-retest coefficients would set an upper bound on correlation between predictor and dependent variables. That it does not probably reflects the fact that conditions of experiments I and II were more similar than those in experiments II and III. The point is that discussions of the validity of CRR should include consideration of the reliabilities of the basic entries into the formula. Considering the stimuli used in this investigation together with the inherent variability associated with responses to these stimuli, CRR did about as well as we could have expected.

Within the limits of this study, CRR was found to be of positive value for the prediction of response patterns to visual form stimuli which were to be presented under a variety of stimulus conditions and viewed by different samples of subjects. The rule accurately predicted exact numeric values when experimental conditions were held constant. When experimental conditions were allowed to vary, the rule did not accurately predict numeric values, but the use of the rule did provide sufficient information to enable the selection of a subset of stimuli from a larger set which sufficiently satisfied pre-established criteria and with far less manipulation of data than other existing techniques.

REFERENCES

Bowen, H. M., Andreassi, J., Truax, S., and Orlansky, J. Optimum symbols for radar display. *Human Factors*, 1960, **2**, 28–33.

Clarke, F. R. The constant-ratio rule for confusion matrices in speech communication. *J. Acoust. Soc. Amer.*, 1957, **29**, 715–720.

Clarke, F. R. Proportion of correct responses as a function of the number of stimulus-response alternatives. *J. Acoust. Soc. Amer.*, 1959, **31**, 835. (Abstract)

Clarke, F. R. and Anderson, C. O. Further test of the constant-ratio rule in speech communication. *J. Acoust. Soc. Amer.*, 1957, **29**, 1318–1320.

Engstrand, R. D. and Moeller, G. The relative legibility of ten simple geometric figures. *Amer. Psychol.*, 1962, **17**, 386. (Abstract)

Hodge, M. H. The constant-ratio rule and identification tasks. USAF: ESD TDR-62-2, January 1962.

Hodge, M. H. and Pollack, I. Confusion matrix analysis of single and multidimensional auditory displays. *J. exp. Psychol.*, 1962, **63**, 129–142.

Hodge, M. H., Crawford, M. J. and Piercy, Mary L. The constant-ratio rule and visual displays. USAF: ESD-TDR-61-56, December 1961.

Hodge, M. H., Piercy, Mary L., and Crawford, M. J. The constant-ratio rule and lifted weights. USAF: ESD-TN-61-61, June 1961.

Pollack, I. and Decker, L. Consonant confusions and the constant-ratio rule. *Lang. Speech.*, 1960, **3**, 1–6.

Shepard, R. N. Analysis of proximities as a technique for the study of information processing in man. *Human Factors*, 1963, **5**, 33–48.