Rated ease of memorization of dot-in-matrix patterns: Multiple regression analysis by physical variables which describe configuration of dots¹

SHIN-ICHI ICHIKAWA²

University of Tokyo³

In Experiment I, 60 8-dot-in- 4×4 -matrix patterns and 60 7-dot-in- 3×5 -matrix patterns were rated for ease of memorization on a 9-point scale. As a result of multiple regression analysis, two effective variables for estimating rated ease of memorization were selected from physical variables which described the configuration of dots in a matrix: One was "equivalence set size" (Garner & Clement, 1963) representing the degree of symmetry, and the other was "concentration along row or column direction" representing the degree of concentration of dots in particular rows or columns. In Experiments IIa and IIb, the performance of immediate recall and that of free recall were shown to be highly correlated with the values estimated by the multiple regression equation which contained the two physical variables as predictor variables. It is concluded that the arrangement of dots can be controlled by the two variables in recall experiments.

Key words: methodology, visual memory, dot patterns, multiple regression analysis, immediate recall, free recall.

In recent years, matrices containing items in their cells have been frequently used as stimuli in studies of memory of spatial information (e.g. Bartram, 1978; den Heyer & Barrett, 1971; Kail & Siegel, 1977; Peterson, 1975). Den Heyer and Barrett (1971) and Ichikawa (1979) showed that a verbal distractor resulted in the loss of letter information while a visual distractor resulted in the loss of position in-

formation. Their results strongly suggest that the positions of items in a matrix are encoded into a form other than that of verbal information. Ichikawa (1979), in addition, showed that visual memory span can be expressed in terms of the number of dots in a 5×5 or 6×6 matrix since the proportion of correctly recalled patterns is a monotonic decreasing function of the number of dots, and he argued that this measure can be employed for comparisons between subjects and between conditions. These studies indicated that this kind of pattern is useful in the investigation of visual memory just as nonsense syllables are in verbal memory. Consequently, it was necessary to find appropriate variables to control the difficulty of memorization caused by the arrangement of items in a matrix, such as association value or meaningfulness for nonsense syllables. This was the purpose of the present research.

In this study, the items in a matrix were all dots in order to simplify the problem. For the arrangement of dots, two different kinds of variables can be employed. One

¹ This study was part of a master's thesis (Ichikawa, 1979), and was reported at the 44th Annual Convention of the Japanese Psychological Association in August, 1980.

² The author is grateful to Prof. Y. Umeoka and Prof. T. Oyama of the University of Tokyo for their many helpful comments. Thanks are due to S. Mayekawa of the University of Tokyo, who devised most of the physical variables used in the present study (Mayekawa, 1975), and helped the author revise the computer program.

³ The author now belongs to Saitama University. Requests for reprints and computer programs should be sent to Shin-ichi Ichikawa, Junior College of Economics, Saitama University, 255 Shimo-ohkubo, Urawa, Saitama 338.

is a psychological variable: for example, rated goodness, rated ease of memorization, and so on. More than likely, these variables will prove to be related to recall performance, as Clement (1967) has shown that rated goodness was highly correlated with performance in paired-associate learning. However, it seems quite laborious to measure such a variable for many patterns. The other sort of variable is a physical one which will be mentioned later in detail. It is defined without human responses, so we can obtain the value of any pattern easily through calculation. Yet, the selection of physical variables affecting recall performance will remain another problem.

One of the resolutions was expected to be multiple regression analysis. In Experiment I, for example, rated ease of memorization of 60 patterns was predicted by physical variables. The variables with high weight were selected, and a multiple regression equation was formed. This equation, however, was originally the one used for predicting ratings. Unless it can estimate recall performance, it has little value for controlling stimulus materials in recall experiments. Therefore, two types of recall experiments were planned in Experiments IIa and IIb to examine whether recall performance would be related to the values estimated by the equation.⁴

Experiment I

Method

Subjects. The subjects were 40 undergraduate students (29 males and 11 females) aged from 18 to 23.

Materials and procedure. Each subject was asked to rate both 60 8-dot-in-4×4-matrix patterns and 60 7-dot-in-3×5-

matrix patterns for ease of memorization. Among these 60 patterns, 40 were produced at random by a computer, but 20 patterns which seemed to be easily memorized were added intentionally. This procedure was adopted because, if all patterns were sampled randomly, only difficult patterns would be selected due to the skewed distribution. The ratings obtained from such random samples would not have much information, for the range of the scale used in the rating would be too small in comparison with the one used in the rating of all possible patterns.

These patterns were numbered from (1) to (60) respectively, and were printed so that 20 patterns were on each sheet of paper; the booklet for one subject had three pages for 60 8-dot-in-4×4-matrix patterns and three pages for 60 7-dot-in-3×5-matrix patterns. The size of a cell in a matrix was 5 mm×5 mm, and the diameter of a dot was 2.5 mm. The subjects were instructed to rate each pattern using a score from 1 to 9 with higher scores representing easier patterns to memorize. Half of the subjects rated 8-dot-in-4×4-matrix patterns first, and 7-dot-in-3×5-matrix patterns later; the other half, vice versa. It took about 15 min for a subject to complete this rating task.

Analysis and Results

The data for 8-dot-in-4×4-matrix patterns and those for 7-dot-in-3×5-matrix patterns received the same analysis separately, so the analysis for the former will be stated in detail first.

The ratings were standardized so that the mean was 0 and the variance was 1 for each subject's scores, because the subjects differed greatly in the means and the variances of their ratings. Thus the ratings were made into a "data matrix", which was 60 (patterns)×40 (subjects) and standardized in each column.

How large is the individual difference in points of view of rating? Is it often found that the same pattern is rated as

⁴ If it were easy to obtain recall performance for each of many patterns, we could directly use it as the criterion variable in the multiple regression analysis instead of rated ease of memorization. The strategy of the present research was adopted because too many recall experiments seemed to be needed for reliable recall data.

"easy to memorize" by one subject but as "difficult to memorize" by another subject? From the data matrix (60×40), a correlation matrix (between subjects, 40×40) was made by computing 780 (i.e. $40 \cdot 39/2$) correlation coefficients. Their distribution was as follows: .1-.2 (1), .2-.3 (8), .3-.4 (31), .4-.5 (79), .5-.6 (126), .6-.7 (216), .7-.8 (273), .8-.9 (46). According to the result of principal component analysis for the correlation matrix, only the first eigenvalue was far larger than the others: The eigenvalues were 26.63, 1.63, 1.38, 1.08, This result suggested that points of view of the 40 subjects were quite similar. In addition, the correlation coefficient between the mean standard ratings of 20 randomly selected subjects and those of the other 20 subjects was .974. Therefore, it seemed reasonable that the mean of the 40 standard ratings for a pattern should be considered as the rated ease of memorization (abbreviated to "REM" in this paper) of the pattern.⁵ The standard deviation of the REM of the 60 patterns was .808.

Multiple regression analysis was carried out to look for important factors which influenced the REM of the patterns. The criterion variable was REM, and the predictor variables were 25 kinds of physical variables which described the configuration of dots in a matrix. For example, a variable DM is the mean distance among dots, SYM5 is the proportion of the number of dots which are put on others when the pattern is rotated 180° , and so on. Some variables are composite functions of the two or more other variables. (See Appendix for detailed definitions of these physical variables.)

The variables with high weight were chosen by the following procedure.⁶

1. First, 25 correlation coefficients between the criterion variable (REM) and successively selected one of predictor variables (25 physical variables) were calculated. The predictor variable giving the highest correlation coefficient was chosen as the first one.

2. The multiple correlation coefficients which were obtained when one of the other 24 predictor variables was adopted as the second, were successively calculated. The predictor variable giving the highest multiple correlation coefficient was chosen as the second one.

3. In like manner, the predictor variables giving the highest multiple correlation coefficient were chosen one after another.

As a result, the following variables proved to be very effective for the prediction of REM.

1. ESS (equivalence set size): The number of patterns resulting from reflection and rotation in 90° -steps. The original pattern itself is counted as one. The more symmetric pattern has the smaller ESS value. This variable was proposed by Garner and Clement (1963). (Mean: 6.57; *SD*: 2.28)

2. CRC (concentration along row or column direction): This is a composite function of CR and CC. CR is the degree of concentration of dots along row direction, and CC is that along column direction (see Appendix). Then,

$$CRC = 1 - (1 - CR)(1 - CC). \quad (1)$$

A large CRC means that dots are more concentrated in particular rows or columns. (Mean: .162; *SD*: .111)

The multiple regression equation was,

$$\begin{aligned} \text{REM}_{\text{est}}^* = & -.692 \text{ ESS}^* \\ & + .385 \text{ CRC}^*. \end{aligned} \quad (2)$$

Here, an asterisk indicates that the values

⁵ We could adopt the first principal component scores as REM of the patterns, but it makes little difference because the correlation coefficient between the mean standard ratings and the first principal component scores is .9998.

⁶ This procedure is equivalent to the so-called "forward selection procedure" in which the variables giving the highest partial correlation coefficient are added one after another (Yanai, 1974).

Table 1

The physical variables selected successively to estimate rated ease of memorization of dot-in-matrix patterns (Experiment 1)

8-dot-in-4×4 patterns		7-dot-in-3×5 patterns	
Physical variable	Multiple correlation coefficient	Physical variable	Multiple correlation coefficient
ESS	.762	ESS	.702
CRC	.850	CRC	.818
CRD	.864	RBP	.850
DOC	.887	SYM2	.866
M2C	.894	SYM1	.888
RBP	.898	M2C	.897
SYM2	.902	DM	.904
SYM6	.906	SYM5	.910
SYM1	.916	CRD	.913
DM	.919	DOC	.916
all (25)	.935	all (23)	.940

of the variable are standardized so that the mean is 0 and the variance is 1, and

REM^*_{est} represents the estimated value (least-square approximation) for REM^* . The multiple correlation coefficient obtained by using this equation was .850. The absolute correlation coefficient between ESS and CRC was very small ($r = -.181$).

The prediction by Equation (2) can be considered to be fairly accurate, but may not be accurate enough in that the proportion of accounted variance (i.e. square of the multiple correlation coefficient) is only 72%. Nevertheless, we would regard the above two physical variables as quite influential for the following three reasons.

1. Addition of the second variable CRC resulted in a rapid increase in the multiple correlation coefficient, whereas addition of successive variables gave only slight increases. Table 1 shows the multiple correlation coefficients obtained when the predictor variables giving the highest coefficient were added one after another.

Fig. 1. Illustration of 8-dot-in-4×4-matrix patterns based on the multiple regression equation (2). ESS and CRC are physical variables which represent the degree of symmetry and the degree of concentration of dots in particular rows or columns, respectively. The projection of each pattern onto the axis REM_{est} indicates its estimated value of rated ease of memorization.

2. A far larger multiple correlation coefficient was acquired when CRC was added as the second predictor variable than when any other variable was adopted as the second. On the other hand, whichever variable might be selected as the third, there was little difference in the resulting multiple correlation coefficients. This was also the case in choosing the successive variables.

3. Only these two variables were selected as the first and the second ones in the analysis of the data for 7-dot-in-3×5-matrix patterns, as is stated later.

Some of 8-dot-in-4×4-matrix patterns used in Experiment I are illustrated in Fig. 1 to represent Equation (2). This equation means that REM^*_{est} is a scalar-product of a vector (ESS^* , CRC^*) and the weight vector $(-.692, .385)$. Accordingly, the distance between the origin and the projection of the point (ESS^* , CRC^*) onto the line with the same direction as the weight vector, is directly proportional to REM^*_{est} or REM_{est} . Each pattern was arranged based on the standardized values, but the scales written

along the axes were converted back to the original values of ESS and CRC.

The same analysis was carried out for the data for 7-dot-in-3×5-matrix patterns. REM of a given pattern was again defined as the mean of the 40 subjects' standard ratings for the pattern; the standard deviation of REM was .809. (The eigenvalues of the 40×40 correlation matrix were 26.78, 1.53, 1.37, 1.12, The correlation coefficient between the mean standard ratings of half of the subjects and those of the other half was .972.) Multiple regression analysis showed that, in this case as well, ESS and CRC were particularly influential (see Table 1). It should be noted that the value of ESS can be only 1, 2, or 4 when a matrix is rectangular, because neither reflection on a diagonal axis nor 90°-rotation is possible. The multiple regression equation was,

$$REM^*_{est} = -.646 ESS^* + .425 CRC^* \quad (3)$$

The multiple correlation coefficient obtained by using this equation was .818, and the correlation coefficient between ESS

Fig. 2. Illustration of 7-dot-in-3×5-matrix patterns based on the multiple regression equation (3).

and CRC was $-.132$. Figure 2 is an illustration of some patterns to represent Equation (3).

Experiment IIa

Experiments IIa and IIb were designed to examine whether the multiple regression equation thus obtained could be generally used for controlling stimulus materials in actual recall experiments. Here, the patterns were so sampled that the values estimated by Equation (2) were high, medium, and low. Two types of recall experiments were conducted to see whether these levels of estimated REM might affect recall performance. The types of experiments were immediate recall after brief exposure in Experiment IIa, and free recall after presentation of a list of patterns in Experiment IIb.

Method

Subjects. The subjects were 24 undergraduate students (12 males and 12 females) aged from 20 to 25.

Materials. Three sets consisting of 10 8-dot-in-4×4-matrix patterns were made according to the values estimated by Equation (2).

- Set High: REM_{est} was from .667 to 1.580
- Set Medium: REM_{est} was from $-.150$ to .175
- Set Low: REM_{est} was from $-.808$ to $-.676$

Here, REM_{est} is REM^*_{est} multiplied by the standard deviation of REM, i.e. .808. All these patterns were different from the ones used in Experiment I, because the purpose of the present experiment was to show that we can generally use Equation (2) in order to control the recall performance of any 8-dot-in-4×4-matrix pattern.

Each pattern was drawn on a white card. The size of a cell was 1 cm×1 cm, and the diameter of a dot was 5 mm.

Procedure. The apparatus was Ralph Gerbrands, Harvard Tachistoscope Mod-

Table 2

Means and *SDs* of the number of immediately recalled patterns as functions of the level of REM estimated by Equation (2)

Level of REM_{est}	Mean	<i>SD</i>
High	8.92	1.38
Medium	6.58	1.72
Low	2.38	2.37

el T-4A. The distance from the subject's eye to the stimulus field was 83 cm, and the luminance of the field was 40 cd/m².

A signal, which was a 4 cm×4 cm square on a white card, was presented for 1 s. One second after the signal disappeared, a stimulus pattern was presented for 100 ms. A subject was required to recall the pattern immediately after the presentation by marking the positions of dots on an answer sheet where 4×4 grids (size of a cell was 5 mm×5 mm) were drawn. If s/he finished recalling the pattern, the next trial began. The presentation order was randomly changed for each subject. Before experimental trials, every subject was informed that all patterns would have 8 dots, and s/he observed five patterns as examples.

Results

Only the perfectly recalled patterns were considered to be correct, since it seemed difficult to give appropriate partial scores for various kinds of errors. Table 2 shows the means and the standard deviations of the number of correctly recalled patterns. There was no case in which the relation between the levels of REM_{est} and the numbers of correct responses was reversed in a given subject, although three subjects had tied scores for Set High and Set Medium. Wilcoxon's signed ranks test showed that both the difference in the scores between Set High and Set Medium, and that between Set Medium and Set Low, were highly significant ($p < .001$ in common). Thus it can be said that Equation (2), which is the one for estimating

REM by ESS and CRC, is useful for controlling the performance of immediate recall after brief exposure.

Experiment IIb

Method

Subjects. The subjects were 24 undergraduate students (12 males and 12 females) aged from 19 to 21. They were all different from those who took part in Experiment IIa. Three groups consisting of eight subjects for each were made.

Materials. The same three sets of patterns used in Experiment IIa were employed. Ten patterns in a set were photographed together; those were arranged in two rows consisting of five patterns each. For one set, four kinds of slides were made by changing the locations of the patterns at random.

Procedure. The subjects in each group were required to learn one set of patterns presented for four trials. In one trial, a set of 10 patterns were projected simultaneously on a screen for 1 min; then, the subjects recalled them for 2 min by marking the positions of dots on an answer sheet where 10 4×4 grids (size of a cell was 5 mm $\times$ 5 mm) were drawn. An experimenter gathered the answer sheets at the

end of a trial and delivered new sheets; the next trial began about 1 min after the end of the preceding duration for recall; another slide for the same set was used for the presentation.

The recall procedure was free recall; the subjects were instructed before the trials that they did not have to worry about the order of recall or the locations of the patterns. They were also informed that only the perfectly recalled patterns would be scored, and were urged to recall all eight dots in a pattern.

Results

Figure 3 shows the mean number of correctly recalled patterns. They were clearly separated by the levels of REM_{est} . These results suggest that Equation (2) can also control the performance of free recall after presentation of a list of patterns.

Discussion

Control of Arrangement of Dots

When dot-in-matrix patterns (or patterns containing other items in matrix cells) are used as materials in memory experiments, the dots are usually arranged in either of two ways to make homogeneous stimuli; one is to arrange dots at random, and the other is to arrange them according to some restrictions. The former may not be preferable if we intend to obtain scores with small variance through a small number of trials, because some patterns extremely easy to memorize, though they are produced at random, will disturb the scores. The latter, on the other hand, is to be preferred as long as the restrictions are shown to be capable of controlling recall performance. The results of Experiments IIa and IIb confirmed the control of dot patterns as stimulus materials by means of ESS and CRC. Even when we use other types of patterns than 8-dot-in- 4×4 -matrix, we will be able to obtain patterns with similar levels of difficulty by choosing patterns whose ESS and CRC are alike, as-

Fig. 3. Mean number of patterns correctly recalled after 1 min presentations of a list of 10 patterns as a function of the number of trials and the level of REM estimated by the multiple regression equation (2).

suming that ESS and CRC are still influential. Of course, it is not certain that any kinds of dot-in-matrix patterns can be controlled well by these two physical variables only. Other variables may be selected when the matrix size or the number of dots contained is quite different.

REM and Rated Goodness of Patterns

Though the REM of dot-in-matrix patterns has never been measured, their goodness has been investigated in perceptual psychology. The ratings obtained in Experiment I suggest that there is a very high correlation between REM and rated goodness. Consequently, it is worthwhile to compare the physical variables affecting REM with the factors found in previous studies on pattern goodness. The same cognitive process may be involved in the two types of ratings.

Garner and Clement (1963) argued that pattern goodness is related inversely to the size of a set of equivalent patterns. In their experiment, the subjects rated 90 5-dot-in-3 $\times$ 3-matrix patterns for goodness. The variable ESS could account for 73.4% of the variance of the ratings. Furthermore, they also suggested that the number of straight lines in a pattern was an influential factor. Their results were supported by Experiment I, where ESS and CRC were selected as important physical variables affecting the REM of patterns. CRC as well as the number of straight lines represents the degree of concentration of dots along row or column direction, but CRC has an advantage as compared with the latter: A line consisting of three or fewer dots in a 4 $\times$ 4 matrix, for example, is ignored if the number of straight lines is employed, but CRC takes this into account.

We should refer to "transformation structure theory" of Imai and his associates (Imai, 1977; Imai, Ito, & Ito, 1976). They argued that pattern goodness is related to the intrapattern transformation structures, which were defined by invariant

properties for the three cognitive transformations, that is, mirroring on a diagonal axis, mirroring on the vertical axis, and 180°-rotation. According to their hypotheses, (a) if a pattern does not have invariant properties for any transformation, its goodness is the lowest, (b) if a pattern is invariant for one transformation, it is rated better than the patterns with no transformation structure; then, the ratings may differ when the transformation structures are not identical, (c) if a pattern is invariant for two transformations, its goodness is better than the one which is invariant for either of the two. Imai et al. (1976) showed that the ratings for 5-dot-in-3 $\times$ 3-matrix patterns supported these hypotheses. On the other hand, Matsuda (1978) pointed out that their theory could not account for the significant differences of ratings of some patterns which had the same transformation structure.

The present paper may provide the methodology to make the prediction of pattern goodness more accurate by considering some variables at once. The author, in fact, carried out multiple regression analysis for Matsuda's (1978) data in order to substantiate the validity of estimating rated goodness by the two variables ESS and CRC. Matsuda required 30 subjects to rate 48 5-dot-in-3 $\times$ 3-matrix patterns for goodness on a 7-point scale, and regarded the mean rating as the rated goodness of a given pattern. The multiple correlation coefficient was .917. This fact means that ESS and CRC are useful for estimating not only REM but also rated goodness of patterns.

References

- Bartram, D. J. 1978 Post-iconic visual storage: Chunking in the reproduction of briefly displayed visual patterns. *Cognitive Psychology*, **10**, 324-355.
- Clement, D. E. 1967 Paired-associate learning as a correlate of pattern goodness. *Journal of Verbal Learning and Verbal Behavior*, **6**, 112-116.
- Garner, W. R., & Clement, D. E. 1963 Goodness

- of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior*, **2**, 446-452.
- den Heyer, K., & Barrett, B. 1971 Selective loss of visual and verbal information in STM by means of visual and verbal interpolated tasks. *Psychonomic Science*, **25**, 100-102.
- Ichikawa, S. 1979 A study of visual memory image by dot-in-matrix patterns. Unpublished master's thesis, University of Tokyo. (In Japanese)
- Imai, S. 1977 Theories of pattern goodness. *Japanese Psychological Review*, **20**, 258-272. (In Japanese)
- Imai, S., Ito, T., & Ito, S. 1976 Effect of intra-pattern transformation structures upon goodness judgments of two dimensional patterns. *Japanese Journal of Psychology*, **47**, 202-210. (In Japanese with English summary)
- Kail, R.V., Jr., & Siegel, A. W. 1977 Sex differences in retention of verbal and spatial characteristics of stimuli. *Journal of Experimental Child Psychology*, **23**, 341-347.
- Matsuda, T. 1978 Judgment of pattern goodness and intraconfigurational transformation structures: An examination of Imai's theory. *Japanese Journal of Psychology*, **49**, 207-214. (In Japanese with English summary)
- Mayekawa, S. 1975 Physical and psychological structures of dot patterns. Unpublished bachelor's thesis, University of Tokyo. (In Japanese)
- Peterson, M. J. 1975 The retention of imagined and seen spatial matrices. *Cognitive Psychology*, **7**, 181-193.
- Yanai, H. 1974 Problems on selection of predictor variables in relation with regression analysis and discrimination analysis. *Japanese Psychological Review*, **17**, 14-28. (In Japanese)
- Young, F. W., de Leeuw, J., & Takane, Y. 1976 Regression with qualitative and quantitative variables: An alternating least squares method with optimal scaling features. *Psychometrika*, **41**, 505-529.
- (Received Sep. 16, 1980; accepted May 16, 1981)

Appendix

The definitions of the physical variables which describe the configuration of dots in a matrix, are as follows. It is supposed that the matrix size is $m(\text{rows}) \times n(\text{columns})$ and the number of dots contained is d . The variables (10), (11), and (13) cannot be defined when a matrix is rectangular (or $m \neq n$).

(1) CR: concentration along row direction.

Let d_i be the number of dots contained in row i ($1 \leq i \leq m$), and let $p_i = d_i/d$. Then,

$$CR = 1 - \left(\sum_{i=1}^m p_i \log_2 \frac{1}{p_i} \right) / \log_2 m,$$

$$\text{If } p_i = 0, \quad p_i \log_2 \frac{1}{p_i} = 0.$$

A large CR means that dots are more concentrated in particular rows. Thus, this variable derives from "redundancy" in the information theory.

(2) CC: concentration along column direction.

(3) CRD: concentration along right-to-left diagonal direction.

(4) CLD: concentration along left-to-right diagonal direction.

(5) CRC: concentration along row or column direction.

$$CRC = 1 - (1 - CR)(1 - CC).$$

(6) CD: concentration along right-to-left or left-to-right diagonal direction.

$$CD = 1 - (1 - CRD)(1 - CLD).$$

(7) CT: total concentration.

$$CT = 1 - (1 - CR)(1 - CC)(1 - CRD)(1 - CLD).$$

(8) SYM1: degree of symmetry, type 1.

through (13) SYM6: degree of symmetry, type 6.

These six variables are proportions of the number of dots which are put on others when the pattern is transformed symmetrically. The transformations are (8) reflection on the vertical axis, (9) reflection on the horizontal axis, (10) reflection on the right-to-left diagonal axis, (11) reflection on the left-to-right diagonal axis, (12) 180°-rotation, (13) 90°-rotation, respectively.

(14) SYMT: total degree of symmetry.

$$\text{If } m = n, \quad SYMT = 1 - (1 - SYM1)(1 - SYM2) \dots (1 - SYM6),$$

$$\text{If } m \neq n, \quad SYMT = 1 - (1 - SYM1)(1 - SYM2)(1 - SYM5).$$

(15) M2R: second moment about row-centroid.

The row-centroid (c_r) is defined by

$$c_r = \sum_{i=1}^m i p_i.$$

Then,

$$M2R = \frac{1}{d} \sum_{i=1}^m p_i (i - c_r)^2.$$

A large M2R means that the distribution of dots is more concentrated near the row-centroid.

(16) M2C: second moment about column-centroid.

(17) M2S: sum of second moment.

$$M2S = M2R + M2C.$$

(18) M3R: absolute value of third moment about row-centroid.

$$M3R = \frac{1}{d} \left| \sum_{i=1}^m p_i (i - c_r)^3 \right|.$$

A large M3R means that the distribution of dots about the row-centroid is more skewed.

(19) M3C: absolute value of third moment about column-centroid.

(20) M3S: sum of third moment.

$$M3S = M3R + M3C.$$

(21) DOC: distance between origin and centroid.

This is the distance between the origin (the center of the matrix) and the centroid (the center of gravity) assuming that the length of a side of a cell is unity.

(22) DM: mean distance.

This variable is the mean distance among dots assuming that the length of a side of a cell is unity.

(23) RBP: redundancy of basic patterns.

Basic patterns mean all the 16 possible 2×2 patterns: one pattern containing no dots, four patterns

containing one dot, six patterns containing two dots, four patterns containing three dots, and one pattern containing four dots. Let the relative frequency with which a basic pattern k ($1 \leq k \leq 16$) appears in the original dot-in-matrix pattern be q_k . Then,

$$RBP = 1 - \left(\sum_{k=1}^{16} q_k \log_2 \frac{1}{q_k} \right) / \log_2 16.$$

(24) SPL: degree of simplicity.

$$SPL = 1 - (1 - CT)(1 - SYMT).$$

(25) ESS: equivalence set size.

This is the number of patterns resulted from transformations of (8), (9), ..., (13), and 270° -rotation. If $m \neq n$, transformations are (8), (9), and (12). The original pattern itself is counted as one.

These physical variables are not always correlated linearly with REM, so the estimation of REM will be more accurate if they receive appropriate monotone transformation. In fact, the author carried out nonmetric multiple regression analysis called MORALS (Young, de Leeuw, & Takane, 1976) using the two variables ESS and CRC, and obtained higher multiple correlation coefficients, i.e. .904 for 8-dot-in- 4×4 -matrix patterns and .844 for 7-dot-in- 3×5 -matrix patterns. However, we do not have a computer program to select influential predictor variables according to the procedure mentioned in page 71.