ON CATEGORICAL SCALES OF UTILITY¹

KAZUO NAKATANI

Senshu University

Utility functions concerning supposed moneys were obtained at the level of interval scale by means of the method of successive categories. They were found to conform to the Fechnerian logarithmic equation, while ratio utility scale constructed by double stimulus method had been found to be congruent with the Stevens' power law. Individual differences in the categorical ratings were discussed by factor analytic technique. A new and simple calculation method based on simplex correlation matrix was proposed on assumption of random walking model about judgment variability.

Mathematical studies on utility began in the 18th century. Since the *expected value* can not explain the gambling behavior, the *moral expectation* defined by *mean utility* became to be used for measurement of risk. The utility had been assumed to be a non-linear function of the value or the quantity of goods. Psychologists and economists of the 19th century discussed the relation between subjective *utility* and objective *value* from view-points of psychophysics or utilitarian philosophy. At first the utility was taken to be approached directly by introspection method, and later to be estimated indirectly by observation of expenditure. But in the early 20th century it became dubious whether the magnitude of utility is scalable in the same sense as the length and the weight of thing. Some of those statisticians, economists and psychologists who were interested in the game theory, in which choice behavior should be performed normatively so as to maximize expected utility, were content with comparisons of utilities at most at the level of hyper-ordinal scale.

Their view-points, methods and theoretical constructs are so divergent that we cannot identify the meanings of the *utility*. The most basic and simplest

definition may be given in terms of subjective value of goods with some unit *utile*. The so called *marginal utility* is no more than the *differential* or *difference* of the subjective value. Although the concept *utility* was a theoretical construct, it turned out to be a basic tool for explaining the most empirical law of Engel.

However, the meanings of the *subjective* are still open to discussion. The divergency of the meanings may be seen to be reflected in the methods. Some psychologists use several kinds of psychophysical judgment methods to scale the *subjective value*, while some economists try to analyse it from micro or macro expenditure data. The difference among methods to approach the utility resulted in the difference among theories concerning the type of utility function. Some insist power law, while others on logarithmic hypothesis. Still some others adopt no specific function.

These differences can be explained in terms of scale levels. Most psychophysicists will measure *cardinal utility*, some at the level of ratio scale and others at the level of interval scale. Many economists who are doubtful about *cardinal utility* define *ordinal utility* for the description of consumer's behavior by *indifference curve*. Instead of utility function, they are likely to use *preference function* or *utility indicator* which are free from the discussion

¹ I am indebted to Dr. Tanaka and Dr. Torii of Tokyo University for aid in preparing this paper.

of the scale levels. But psychologists are interested in utility metrics because it will contribute to the issues developed on sensory scalings and human numerical rating behaviors.

It was under the influence of Bernoulli's paper (1738) that Fechner proposed the logarithmic equation between stimulus magnitude and sensation magnitude. Since then the law was the greatest target of psychophysical issues. The most powerful counter proposal was made by Stevens that psychological scale constructed by direct methods is power function of physical scale. One can find an archetypal divergence between Bernoulli's and Cramer's views about utility function. Following Cramer's opinion of 1728, law of diminishing marginal utility can be expressed by differential equation:

$$\frac{dy}{dx} = \frac{1}{y}$$

where y =utility, dependent variable whose unit may be taken appropriately
 x =quantity of goods, independent variable,

while Bernoulli assumed it to be

$$\frac{dy}{dx} = \frac{1}{x}.$$

They defined *diminishing* by *inversely proportionate* to dependent or independent variable. Cramer's function:

$$y = \sqrt{x}$$

which is a solution of the former differential equation is a specific case of Stevens' power law with characteristic exponent 0.5. Bernoulli's function:

$$y = \log x + c$$

which is a solution of the latter is the classical model of Fechner's logarithmic law (Piéron 1963).

Stevens (1962, 1968) supports Cramer's proposal, citing Galanter's experiment

and Selling and Wolfgang's research (1964). We may assert that if we adopt *direct* methods the subjective scales of utility will be constructed at the level of ratio scale and satisfy the power law. Although subjects will be embarrassed when asked to estimate subjective magnitudes of numerical stimuli such as money and quantity of goods by *direct* methods, use of some additional questions or modifications of the methods will make it easy to produce numerical judgments. Our scaling experiment by double stimulus method showed the characteristic exponent near to 0.63 (Nakatani, 1968b).

Fechnerian utility scales have been constructed by *indirect* methods such as successive categories method at the level of interval scale. As in the general sensation metrics, the categorical utility scales are said to be in non-linear, i.e., concave upward relationship to the ratio utility scales.

Two points should be checked in order to make it clear why discrepancy appears between the ratio scales and the category scales. The one is on *direct* methods, that is, whether the ratio values computed from subjects' numerical responses satisfy ratio axioms to deserve the name of ratio scale (Nakatani, 1967). The other is on validity of basic assumptions or principles of *indirect* methods. The present paper is especially concerned with the latter problem and presents a new and simple calculation method for the categorical methods.

PROCEDURE

The subjects were 43 students of Chuo Denki Gakuin, taking course on general psychology. They were asked to rate *utility* of money by method of successive categories with the following questionnaire:

"Suppose there is such an amount of money as follows (from A to O) on your hand, how do you feel it, large or small? Please rate your feeling about each money in accordance with

the 7-point scale:

code	successive categories
1	extremely large
2	considerably large
3	rather large
4	medium
5	rather small
6	considerably small
7	extremely small

and write down the appropriate coding number in the following right column.

	amount of money	answer
A	10,000,000. yen	
B	5,000,000. yen	
C	1,000,000. yen	
D	500,000. yen	
E	100,000. yen	
F	50,000. yen	
G	10,000. yen	
H	5,000. yen	
I	1,000. yen	
J	500. yen	
K	100. yen	
L	50. yen	
M	10. yen	
N	5. yen	
O	1. yen	

RESULTS

Three kinds of cardinal utility scales can be constructed by application of psychophysical scaling methods to ordinal data. They are mean category scale, distribution category scale and simplex category scale. The last is calculated with simplex correlation matrix on the assumption of random walk model about variability of categorical ratings.

I. Mean Category Scale

Values of this scale can be obtained as simple arithmetical means of rating data

TABLE 1
Frequency of rated successive categories

<i>S</i> (yen)	Rating code						
	7	6	5	4	3	2	1
1×10^0	41	2					
5×10^0	39	3	1				
1×10^1	34	8	1				
5×10^1	28	12	3				
1×10^2	22	9	11	1			
5×10^2	10	14	15	4			
1×10^3	4	13	16	9	1		
5×10^3	1	7	17	16	1	1	
1×10^4		5	12	18	7	1	
5×10^4		2	7	18	9	6	1
1×10^5		1	3	15	12	11	1
5×10^5			2	6	16	11	8
1×10^6			2	2	10	17	12
5×10^6				2	2	17	22
1×10^7					1	8	34

by weighting some successive integers to the successive categories. It seems to be most expedient technique; however it might be criticized that the distances between adjacent categories are treated as equal.

Table 1 shows frequencies of the categories that the 43 subjects assigned to the 15 supposed moneys. According to general usage, the assigned category will be designated by *R* and the supposed money by *S*. Fluctuations of the *R*s on the middle *S*s are approximately normal, which suggest that the calculations of arithmetical mean on the assumption of equally spaced categories seem reasonable.

It should be noted that the mean category scale is an interval scale because the boundary numbers of the successive integers are arbitrary, that is, the category weights may be either (+3, +2, +1, 0, -1, -2, -3) or (7, 6, 5, 4, 3, 2, 1, 0) and so on. We use provisionally the first case of the weights. If necessary, we can change its origin and unit by some appropriate linear transformation.

Fig. 1 shows the mean values, designated

FIG. 1. Mean category scale of utility.

R_i , plotted against $\log S$. For the stimuli, greater than 10 yen, this semi-log graph fits fairly well to the Fechnerian equation:

$$R_i = a \log S + b.$$

For the smaller stimuli the points tend to deviate from the equation. Such tendencies are often reported in the categorical rating data of subjective sensation magnitudes about weightness, length and so on. When the same questionnaire was applied to about thousand housewives living in up- and down-town of Tokyo, mean values of utility ratings showed very similar curve. The only difference is that in the domain more than 500 yen a logarithmic equation can be fitted and linear equation in the smaller domain. It might be possible that utility function for the accustomed amount of money is linear. But we cannot also neglect the possibility that the scale values are biased by the application of the mean category method itself (Nakatani, 1968a).

As the fluctuation of the rating cannot overflow beyond the extreme categories, variances of the ratings about four stimuli at the both extremes (A, B, N and O) are relatively small (Fig. 2). There is no reason to assume that both the outer and the inner categories are equally spaced. The tails of the curve may be distorted by the successive integer weightings. Of course the distortions

FIG. 2. Standard deviation of the ratings.

may exist even in the case of the middle stimuli; they, if any, are not so large as in the case of the extreme stimuli.

II. Distribution Category Scale

Without weighting *a priori* integers to the categories we can construct scales on Thurstonian model about the discriminational process named *law of categorical judgment* (Torgerson, 1958). Both the subjective magnitude R and the categorical boundary t are postulated to fluctuate along latent psychological continuum by the normal distribution function. According to this law,

$$t_g - R_i = Z_{ig}s_{ig}$$

where t_g = mean location of the g -th category boundary

R_i = scale value of stimulus i

Z_{ig} = unit normal deviate corresponding to the proportion of times S_i is sorted below g -th categorical boundary

s_{ig} = standard deviation of the difference between subjective magnitude i and categorical boundary g .

Table 2 shows the Z scores transformed from Table 1. There are many unfilled cells in the north-east and south-west parts of the table. The Z scores corresponding to the cumulative proportions near 0% or 100% are useless,

TABLE 2
Matrix of Z_{ig}

S_j		Categorical boundaries t_g					
j	(yen)	t_1	t_2	t_3	t_4	t_5	t_6
1	1×10^0	1.68					
2	5×10^0	1.32	2.00				
3	1×10^1	.81	2.00				
4	5×10^1	.39	1.48				
5	1×10^2	.03	.59	2.00			
6	5×10^2	-.73	.15	1.32			
7	1×10^3	-1.32	-.27	.73	2.00		
8	5×10^3	-2.00	-.89	.20	1.68	2.00	
9	1×10^4		-1.20	-.27	.89	2.00	
10	5×10^4		-1.68	-.81	.33	.98	2.00
11	1×10^5		-2.00	-1.32	-.15	.59	2.00
12	5×10^5			-1.68	-.89	.15	.89
13	1×10^6			-1.68	-1.32	-.45	.59
14	5×10^6				-1.68	-1.32	-.03
15	1×10^7					-2.00	-.81

FIG. 3. Linear relations between adjacent Z scores, Z_{ig} and $Z_{i+1,g}$.

not only because of low reliability but also of bias owing to the non-linear transformation. In order to solve the unknown R from such a Z matrix, it is most convenient and practical to use graphical method with additional assumption:

$$s_{ig} = 1.$$

Then taking t as parameter, linear relations between adjacent rows of the Z score matrix will be obtained. That is,

$$Z_{ig} = Z_{i+1,g} + (R_{i+1} - R_i).$$

Plotting the relations between Z_{ig} and $Z_{i+1,g}$ in Fig. 3, the difference $R_{i+1} - R_i$ is observed as the ordinate intersection. Even when there is only one point, we can estimate the intersection because the number of unknown parameter is only one. When there are more than two points, the estimation should be determined with the point that is the nearest to the intersecting point. The direction coefficient of the line should be always one, even if any other regression line seems to fit more fairly.

The scale values given by this equation, say R_{II} , is also of interval scale. Taking the value of the least stimulus S_1 as a

FIG. 4. Distribution category scale of utility.

provisional origin, we can determine the other values sequentially. Thus, calculated values are shown in Fig. 4 which satisfy the Fechnerian hypothesis approximately.

Although the calculation technique is very simple, there are lots of disadvantages in this scaling procedure. For example, we neglect large amount of informations contained in the raw data by using only a point in order to estimate the intersection. The most intrinsic defect is that relations among subjective ratings are not considered. Each rating is treated as if it fluctuates independently to another. Practically, however, the assignment of a category to a stimulus is always performed in the whole context of the relations among the categories and the stimuli. Furthermore individual differences can not be neglected. A same category may be different in width from one subject to another. Each subject may have his own manner of categorical judgment: for some subjects the upper categories are wide and for others narrow. Variability of categorical rating is better evaluated in terms of the individual difference of the manner of rating.

III. Simplex Category Scale

When the stimuli are presented sequentially in some order as in our case, it might be able to know some information about the manner of rating by examining the length of run of each category. But more general and refined ways to find the structure of the rating manner may be available by factor analysis of the correlation matrix among stimulus ratings. Since our raw data are on ordinal scale only, we cannot start our discussion from the cross-product correlation coefficients in rigorous sense, however, the assignments of some successive integers to the successive categories as in the case of mean category scaling make it possible to introduce most convenient approximation to the correlation matrix (Table 3).

TABLE 3
Matrix of correlation (r_{ij})

	1×10^0	5×10^0	1×10^1	5×10^1	1×10^2	5×10^2	1×10^3	5×10^3	1×10^4	5×10^4	1×10^5	5×10^5	1×10^6	5×10^6	1×10^7
1×10^0	1.000	.791	.590	.386	.418	.285	.283	.292	.186	.238	.158	.124	.146	.036	-.125
5×10^0		1.000	.741	.477	.536	.356	.386	.365	.285	.297	.243	.113	.170	.166	.021
1×10^1			1.000	.698	.765	.580	.629	.542	.413	.440	.396	.274	.229	.207	.137
5×10^1				1.000	.864	.711	.672	.650	.491	.391	.412	.274	.266	.304	.252
1×10^2					1.000	.852	.828	.739	.602	.480	.503	.323	.330	.342	.212
5×10^2						1.000	.861	.824	.734	.588	.599	.506	.468	.456	.265
1×10^3							1.000	.860	.787	.654	.673	.492	.469	.458	.341
5×10^3								1.000	.836	.711	.739	.635	.611	.568	.414
1×10^4									1.000	.877	.822	.748	.671	.622	.410
5×10^4										1.000	.877	.869	.814	.677	.442
1×10^5											1.000	.881	.802	.708	.437
5×10^5												1.000	.893	.765	.451
1×10^6													1.000	.861	.619
5×10^6														1.000	.738
1×10^7															1.000

The largest correlation in each row is in every case next to the main diagonal. The more distant a cell is from the diagonal, the smaller is the size of the correlation. This decreasing tendency is not necessarily regular. There appears a small negative value in the most distant cell. These disfigurations may be attributable to the small sample size and the integer assignments. They are likely to injure the correlation cells of extreme stimuli. If it were infallible, this correlation matrix might be a perfect simplex.

Calculation were done twice: the first time by using the whole 15×15 matrix and the second time by using the matrix excluding the extreme 4 stimuli (A, B, N and O). Three factors were extracted by means of Horst's successive residual method, which finish the extraction when the eigenvalue becomes less than one (Horst, 1965).

Table 4 shows the eigenvalues and the corresponding factor loadings. The first factor loadings are nearly equal over the all stimuli. The second loadings change their values sequentially from positive to negative. The third change

TABLE 4
Factor analysis

	a_1	a_2	a_3	a_1	a_2	a_3
1×10^0	.377	.622	.580			
5×10^0	.482	.632	.498			
1×10^1	.659	.591	.125	.650	.500	.491
5×10^1	.697	.457	-.268	.706	.553	.140
1×10^2	.789	.470	-.280	.806	.533	.017
5×10^2	.836	.211	-.348	.864	.284	-.250
1×10^3	.864	.188	-.312	.888	.249	-.226
5×10^3	.901	.041	-.197	.914	.067	-.215
1×10^4	.884	-.166	-.100	.899	-.193	-.216
5×10^4	.864	-.276	.174	.863	-.389	.113
1×10^5	.861	-.309	.084	.864	-.377	.064
5×10^5	.779	-.477	.209	.772	-.566	.145
1×10^6	.767	-.501	.269	.732	-.553	.151
5×10^6	.724	-.496	.154			
1×10^7	.507	-.470	-.021			
E. Value	8.421	2.793	1.195	7.372	1.945	.528

their signs twice, from positive to negative and again to positive. This characteristics is somewhat reminiscent of principal components of perfect scale (Guttman, 1950). In the present case, these factors

TABLE 5
Supposed factor analysis (simplified case)

	a_1	a_2	a_3
S_1	1.	1.	.6
S_2	1.	.5	-.3
S_3	1.	.0	-.6
S_4	1.	-.5	-.3
S_5	1.	-1.	.6
eigen value	5.	2.5	1.26

may be called mean component, variance component and skewness component.

The reason why we name in this way is as follow. Suppose, for simplicity, we have rating X_{ij} ($j=1, 2, \dots, 5$) about stimulus j of subject i , and its factor loadings as shown in Table 5.

If a subject, whose three factor scores are all zero, rated as Table 6 (a), his rating

TABLE 6a
Standard rating (factor score F and
reference rating $\bar{x}$)

F_1	F_2	F_3	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$
0	0	0	-2	-1	0	1	2

TABLE 6b
Difference of mean component

F_1	F_2	F_3	x_1	x_2	x_3	x_4	x_5
+1	0	0	-1	0	1	2	3
-1	0	0	-3	-2	-1	0	1

TABLE 6c
Difference of variance component

F_1	F_2	F_3	x_1	x_2	x_3	x_4	x_5
0	+1	0	-1	-.5	0	.5	1
0	-1	0	-3	-1.5	0	1.5	3

TABLE 6d
Difference of skewness component

F_1	F_2	F_3	x_1	x_2	x_3	x_4	x_5
0	0	+1	-1.4	-1.3	-.6	.7	2.6
0	0	-1	-2.6	-.7	.6	1.3	1.4

scores might be considered to be a reference one and can be compared with each subject's rating. This set of x scores is to be obtained as mean category scale before calculating simplex correlation matrix. Linear model under the factor analysis can be written as follows, where $\bar{x}$ is the reference rating:

$$x_{ij} = \bar{x}_j + a_{j1}F_{1i} + a_{j2}F_{2i} + a_{j3}F_{3i} \\ (j = 1, 2, \dots, 5).$$

The ratings of those whose first factor scores are either +1 or -1 and other scores are zero should be as Table 6 (b). The difference between these manners of rating are no more than that of the means.

In a similar way the ratings of those whose second factor scores are either +1 or -1 and other scores zero should be as Table 6 (c). This difference is that of the variance.

And the difference between the ratings of those whose third scores are either +1 or -1 and other scores zero as shown in Table 6 (d) is that of the skewness.

Thus the individual difference of the manner of rating could be represented simply by three orthogonal components. It may be the most essential character of the simplex ratings that the mean, the variance and the skewness component are orthogonal to each other.

Fig. 5 shows curvilinear relation between the first and the second factor loadings, which suggests that utility is not uni-dimensional. However, when we exclude two extreme stimuli on both side i.e., A, B, N and O, we may take it uni-dimensional.

Some scaling methods have been developed in terms of the simplex structure. Since the categorical ratings have the above structure, the psychophysical scaling procedures should also reflect the statistical characteristics. Using the concept of partial correlation, Guttman (1954) made it clear that in perfect simplex correlation matrix $\{r_{ij}\}$ there exist

FIG. 5. Relation between the first and the second factor loadings.

the relation:

$$r_{ij} = \frac{A_i}{A_j} \quad (i \leq j)$$

where A is some parameter named *simplex loading*.

Kaiser (1962) showed a least square solution of the simplex loading A by logarithmic transformation:

$$\log r_{ij} = \log A_i - \log A_j \quad (i \leq j).$$

Schönemann (1970) appreciated this equation not only because it is a tool for the solution but also because it satisfies the distance axioms. We would like to clarify the meaning of the equation on the assumption that the individual manner of rating, i.e., the total feature of over- and under-rating of stimuli, is a randomly walking process along some psychophysical continuum.

It may be well said that the relation between the correlation matrix of finite number of variables and its factor analysis is quite analogous to the relation between the auto-correlation function and its Fourier analysis. If we were able to gather data of continuous ratings along the stimulus continuum, the correlation matrix should take the form of auto-correlation function. The relative size of eigenvalues and the wave forms of the factor loadings of the simplex correlation matrix suggest us that the simplex correlation matrix corresponds to the auto-correlation function of stationary normal stochastic process named first order Markoff series. If such a process is observed on discrete times, $t_i = i$ and $t_j = j$, the correlation can be presented by the form of correlogram:

$$r_{ij} = r(i-j) = r^{j-i} \quad (i \leq j).$$

There underlies first order auto-regression model (Kendall & Stuart, 1966), which can be expressed by the following equations:

$$u_j = r_{ij}u_i + e$$

$$u_{i+1} = ru_i + e$$

where variance of the variate u is one. As it is stationary, the auto-correlation function can be determined only by the distance between two time points:

$$\tau = t_j - t_i,$$

giving the general form:

$$r(\tau) = \exp(-\beta\tau).$$

When the τ is unknown, taking negative logarithm:

$$\tau = -\frac{1}{\beta} \log r(\tau),$$

we can estimate it as least square solution. Thus the distance to be found in time series is equivalent to Kaiser's negative logarithm to be calculated in test series.

Such an exponentially decreasing tendency also exists in our rating correla-

FIG. 6. Decreasing tendency of simplex correlation.

tion matrix. Fig. 6 shows how the correlation decreases as the distance from the diagonal increases. Excluding the correlation coefficients of the extreme four stimuli of low reliability, we find a fairly well fitting regression lines as in Fig. 7 by the negative logarithmic transformation, which is a sufficient evidence of the fact that the stochastic process of the fluctuations of over- and under-ratings is stationary first order Markoffian.

As in the case of the stationary time series, this simplex correlation of the

FIG. 7. Negative logarithmic transformation of the correlation.

categorical rating can be defined by the distance between stimuli;

$$u(S_j) - \bar{u}(S_j) = r_{ij} \{u(S_i) - \bar{u}(S_i)\} + e,$$

$$r_{ij} = \exp \{-d(S_i, S_j)\}$$

where d is a distance function between S_i and S_j .

Let $d(S_i, S_j) = R_j - R_i$. Then for $R_i \leq R_j$,

$$R_j - R_i = -\log r_{ij}.$$

As auto-correlation is even function, for $R_i \geq R_j$,

TABLE 7
Skew symmetrical matrix ($\mp \log r_{ij}$)

	1×10^1	5×10^1	1×10^2	5×10^2	1×10^3	5×10^3	1×10^4	5×10^4	1×10^5	5×10^5	1×10^6
1×10^1	.000	.156	.117	.237	.201	.266	.384	.356	.402	.562	.640
5×10^1	-.156	.000	.064	.147	.172	.187	.309	.408	.385	.562	.576
1×10^2	-.117	-.064	.000	.070	.082	.131	.221	.319	.298	.491	.482
5×10^2	-.237	-.147	-.070	.000	.065	.084	.134	.231	.222	.296	.330
1×10^3	-.201	-.172	-.082	-.065	.000	.066	.104	.184	.172	.308	.329
5×10^3	-.266	-.187	-.131	-.084	-.066	.000	.078	.148	.131	.197	.214
1×10^4	-.384	-.309	-.221	-.134	-.104	-.078	.000	.057	.085	.126	.174
5×10^4	-.356	-.408	-.319	-.231	-.184	-.148	-.057	.000	.057	.061	.090
1×10^5	-.402	-.385	-.298	-.222	-.172	-.131	-.085	-.057	.000	.055	.096
5×10^5	-.562	-.562	-.491	-.296	-.308	-.197	-.126	-.061	-.055	.000	.049
1×10^6	-.640	-.576	-.482	-.330	-.329	-.214	-.174	-.090	-.096	-.049	.000
Mean	-.302	-.241	-.174	-.083	-.058	-.003	+.072	+.136	+.146	+.237	+.271

FIG. 8. Simplex category scale.

$$R_j - R_i = \log r_{ij}.$$

So the estimation of R_j will be given by the column mean of the skew symmetric matrix $\{\mp \log r_{ij}\}$ which is obtained by logarithmic transformation of simplex correlation matrix with positive sign for upper triangle and negative sign for lower triangle (Table 7). The j -th column vector of the matrix is, if infallible,

$$\begin{bmatrix} R_j - R_1 \\ R_j - R_2 \\ \vdots \\ R_j - R_n \end{bmatrix}.$$

The values of R is of interval scale, whose origin is arbitrarily determined to be

$$\sum_i R_i = 0.$$

Estimates obtained in such a way are named *simplex category scale* R_{III} , and are shown in Fig. 8. The abscissa is $\log S$. This utility also fits to the Fechnerian equation fairly well.

DISCUSSION

As the general psychophysical scalings of subjective sensation magnitudes do, it turned out to be true that the subjective

utility of money scaled by categorical method satisfies Fechnerian logarithmic equation. But if we examine the values in detail, we may find systematic deviations from the regression curve. Let's divide the stimuli into two groups, one is the stimuli to be expressed by " 1×10^m ($m=\text{integer}$)" such as 10, 100, 10,000 and so on, and the other by " 5×10^m ($m=\text{integer}$)" such as 50, 500, 50,000 and so on. Then a level shift between the utility functions of the former and the latter is observed. It means that the Fechnerian law does not necessarily fit to such a narrow partial range as "from 1×10^m to $1 \times 10^{m+1}$ ($m=1, 2, 3, \dots$)". Linear equation seems to fit to each interval.

These violations may due to the rating conditions. It can be readily recognized that as the orderings or steps of stimulus presentation are changed, the scale values may be shifted to some extent. And also the alternations of reporting codes of successive categories may possibly influence the values of utilities. Fechnerian equation will be proven only under some optimal conditions about the prepared stimuli and the permitted categories.

We can not say to what extent the utility function obtained in the present experiment concerning supposed money will explain or predict behaviors of those in practical situations of gamble or game. Measurements at the level of ordinal or hyper-ordinal scale might be sufficient for those cases. Some economists are interested more in ordinal utility than in cardinal utility. However, experimental studies of cardinal utility will provide us a basic knowledge about human numerical behaviors in general. It becomes more and more necessary to take verbal reports of numerical categories or codes as raw data in almost all fields: for example, in field researches of social attitudes and in paper tests of personality differences. Without understanding the characteristics of numerical behaviors we cannot adopt correct calculation methods for the data.

The *numbers* reported verbally by subjects are often treated as the same sort of data as numbers read from metrical instruments. Although the *numbers* usually pretend to be *real*, they are often far from satisfying the distance axioms and the ratio axioms (Tanaka & Nakatani, 1967; Nakatani 1967). Subjects can merely match the sensation magnitudes with their some subjective number systems (Ekman, 1964). Relations between ratio scales constructed by *direct* methods and category scales by *indirect* methods can be approached only under the true appreciation of the subjective number behaviors.

REFERENCES

- BERNOULLI, D. 1738 Exposition of a new theory on the measurement of risk. Translated from Latin into English by L. Sommer 1954 in *Econometrica* **22**, 23-36.
- EKMAN, G. 1964 Is the power law a special case of Fechner's Law? *Percept. Mot. Skills* **19**, 730.
- GUTTMAN, L. 1950 The principal components of scale analysis. In S. A. Stouffer *et al*, *Measurement and prediction*. New York: Wiley.
- GUTTMAN, L. 1954 A new approach to factor analysis: The radex. In P. F. Lazarsfeld (ed.) *Mathematical thinking in the social sciences*. Ill: The Free Press, 258-348.
- HORST, P. 1965 *Factor Analysis of Data Matrices*. New York: Holt.
- KAISER, H. F. 1962 Scaling a simplex. *Psychometrika*, **27**, 155-162.
- KENDALL, M. G., & STUART, A. 1966 *The advanced theory of statistics*, vol. 3. London: Griffin.
- NAKATANI, K. 1967 On ratio scales of weight. *Jap. psychol. Res.*, **9**, 155-165.
- NAKATANI, K. 1968a Quantification of utility, I. *Bull. Nikkei Adv. Res. Inst.*, **2** (6), 16-21. (in Japanese).
- NAKATANI, K. 1968b Quantification of utility, II. *Bull. Nikkei Adv. Res. Inst.*, **2** (7), 33-38. (in Japanese).
- PIÉRON, H. 1963 La psychophysique. In P. Fraisse *et al* (eds.) *Traité de psychologie expérimentale*, II. Paris: Presses Univ. de France.
- SCHÖNEMANN, P. H. 1970 Fitting a simplex symmetrically. *Psychometrika*, **35**, 1-21.
- SELLING, T., & WOLFGANG, M. E. 1964 *The Measurement of Delinquency*. New York: Wiley.
- STEVENS, S. S. 1962 Measurement, psychophysics and utility. In C. W. Churchman *et al* (eds.) *Measurement: Definitions and theories*. New York: Wiley.
- STEVENS, S. S. 1968 Ratio scales of opinion. In D. K. Whitla (ed.) *Handbook of measurement and assessment in behavioral sciences*. Massachusetts: Addison-wesley.
- TANAKA, Y., & NAKATANI, K. 1967 On categorical scales of weight. *Percept. Motor Skills*, **24**, 143-150.
- TORGERSON, W. S. 1958 *Theory and methods of scaling*. New York: Wiley.

(Received July 26, 1971)