Perception & Psychophysics
1985, 37 (1), 89-91

Luce's choice model and Thurstone's categorical judgment model compared: Kornbrot's data revisited

ROBERT M. NOSOFSKY
Indiana University, Bloomington, Indiana

Kornbrot (1978) compared Luce's (1963) choice model and the (constant-variance) Thurstone categorical judgment model (Thurstone, 1927; Torgerson, 1958) on their ability to account for unidimensional absolute identification performance. Two subjects were required to identify eight auditory stimuli varying in loudness in a neutral and a payoff-biased condition. Kornbrot concluded that the experimental results provided strong evidence against the choice model of category judgment, and that the Thurstone model yielded impressive fits to the identification data. In this note, I suggest that Kornbrot's conclusion regarding the choice model was too strong. Kornbrot fitted a special case of Luce's choice model to the identification data by constraining the similarity parameters in the general model. The special-case choice model may have provided poor fits because inappropriate constraints were assumed. A choice model with slightly modified similarity constraints yields impressive fits to the identification data, ones that compare favorably with the fits of the Thurstone model.

Assume n stimuli in an absolute identification experiment. According to Luce's (1963) choice model, the probability that a subject makes response j given stimulus i, $P(R_j \mid S_i)$, is given by

$$P(R_j \mid S_i) = \frac{\beta_j \eta_{ij}}{\sum_k \beta_k \eta_{ik}}, \qquad (1)$$

where $0 \leq \beta_j, \eta_{ij} \leq 1$, $\Sigma_{j=1}^{n} \beta_j = 1$, $\eta_{ji} = \eta_{ij}$, and $\eta_{ii} = 1$. The β_j parameters are interpreted as response bias parameters, and the η_{ij} parameters are interpreted as similarity measures on the stimuli S_i and S_j. The reader may verify that there are $n-1$ freely varying response bias parameters and $n(n-1)/2$ freely varying stimulus similarity parameters (one for each pair of distinct stimuli). A constant-variance Thurstone model uses $n-1$ freely varying response bias parameters and $n-1$ freely varying stimulus scale parameters. To equalize the number of stimulus parameters used by the Thurstone model and the choice model, Kornbrot (1978, p. 196, Equation 18) introduced the following constraint equation for the choice model similarity parameters:

$$\eta_{ij} = \eta_{i,i+1}, \; \eta_{i+1,i+2} \cdots \eta_{j-1,j}. \qquad (2)$$

In this equation, the stimuli have been indexed in order of increasing magnitude on the unidimensional continuum. Thus, the similarity between any two stimuli S_i and S_j $(i<j)$ is given by the product of similarities of the adjacent stimulus pairs that link S_i and S_j. This results in $n-1$ free similarity parameters, one for each pair of adjacent stimuli. The use of this constraint equation was motivated, in part, by a suggestion made by Luce (1963, p. 114, Axiom 3).

In the original formulation of the choice model, Shepard (1957, pp. 333, 337) suggested an interpretation of the η_{ij} parameters in terms of distances in a psychological space. He assumed that

$$\eta_{ij} = f(d_{ij}), \qquad (3)$$

where f is some monotonically decreasing function and the d_{ij}s are distances that satisfy the metric axioms (minimality, symmetry, and the triangle inequality). To reduce the number of parameters that needed to be estimated, Shepard suggested that the stimuli be represented as points in a multidimensional psychological space. The d_{ij}s could then be derived by computing the distances between the points in the space. The configuration of points that achieved the best account of the identification data was taken as the multidimensional scaling (MDS) solution for the stimulus set. I will refer to Equation 1 with the assumption that the similarity parameters are functionally related to distances in a psychological space as the MDS-choice model.

Kornbrot's constraint equation (Equation 2) has a simple interpretation in terms of the multidimensional scaling approach to modeling similarity. Assume that the n stimuli can be represented as points on a unidimensional psychological continuum and that distances on this continuum are additive. That is, assume

$$d_{ij} = d_{i,i+1} + d_{i+1,i+2} + \cdots + d_{j-1,j}, \qquad (4)$$

where d_{ij} is the distance between the points representing S_i and S_j. (Again, the stimuli are indexed in order of increasing magnitude on the psychological continuum.) Then the product rule introduced by Kornbrot follows if

$$\eta_{ij} = e^{-d_{ij}}, \qquad (5)$$

Portions of this article appeared in an appendix of the PhD dissertation submitted by this author to Harvard University. The work was supported by Grants BNS 80-26656 from the National Science Foundation and MH 37208 from the National Institute of Mental Health to Harvard University, and by BRSG Grant S07 RR07031 from the Biomedical Research Support Grant Program to Indiana University. I thank Diana Kornbrot and Michael Birnbaum for their helpful comments and suggestions for future research.

The author's mailing address is: Department of Psychology, Indiana University, Bloomington, IN 47405.

Table 1
Chi-Square Values for Best Fitting Thurstone and Choice Models

	Thurstone		Choice		
Subject	Logistic Distribution	Gaussian Distribution	Exponential Decay Similarity Function	Gaussian Similarity Function	df
			Neutral Condition		
1	36.7	40.4	56.9*	35.0	(30)
2	29.1	31.6	52.6*	26.9	(21)
			Payoff Biased Condition		
1	21.9	32.1	289.0*	32.4	(23)
2	19.6	15.5	245.5*	13.1	(20)

*Measured $\chi^2 > \chi^2 5\%$.

that is, if similarity is an exponential decay function of psychological distance.

In previous tests of the MDS-choice model, Shepard (1958a) found support for the assumption of an exponential decay function relating stimulus similarity to psychological distance. He noted in subsequent work (Shepard, 1958b, 1962), however, that the best fitting similarity function might vary as a function of experimental conditions. Recently, Nosofsky (1984) conducted experiments in which subjects identified two-dimensional stimuli. The MDS-choice model was used to account for the confusion data in these experiments. It was found that a Gaussian similarity function provided much better accounts of the data than did an exponential decay similarity function. The use of a Gaussian function was suggested by previous theoretical work conducted by Shepard (1958b), and by the well-known success of the Gaussian distribution in Thurstonian modeling. The Gaussian similarity function takes the form:

$$\eta_{ij} = e^{-d_{ij}^2}. \tag{6}$$

The MDS-choice model studied by Kornbrot may have provided poor fits to her data because an inappropriate similarity function was used. It seemed reasonable to refit the model to her data using a Gaussian similarity function. To apply the model, one simply represents each stimulus as a point on a unidimensional psychological continuum. Letting x_i be the value of the point representing stimulus S_i, the distance between S_i and S_j is given by $d_{ij} = |x_i - x_j|$. Applying Equation 6 and substituting the derived η_{ij} values into Equation 1 generates the theoretical predictions. The parameters in the model are the locations of the points on the unidimensional continuum ($x_1, \ldots, x_n$) and a set of response bias parameters ($b_1, \ldots, b_n$)[1]

Table 1 presents the minimum chi-square values for the various models fitted to Kornbrot's data.[2] The first three columns give the chi-square values reported previously by Kornbrot (1978, p. 206, Table 5). Columns 1 and 2 give the chi-square values for the Thurstone models. (Kornbrot fitted two Thurstone models to the data, one that assumed logistic distributions and the other, Gaussian distributions.) Column 3 gives the chi-square values for the exponential MDS-choice model, that is, Luce's choice model constrained by Equation 2. Column 4 gives the newly computed chi-square values for the Gaussian MDS-choice model.[3] The best fitting parameters for this model are presented in Table 2. The Gaussian similarity function provides far better fits than the exponential decay similarity function. In fact, with the Gaussian assumption, the MDS-choice model provides slightly better fits than either Thurstone model in three of four cases, although it is slightly worse on the fourth.

A graphical comparison of the Gaussian and exponential decay similarity functions is provided in Figure 1 for Subject 1 in the payoff-biased condition. The following steps were followed in developing these graphs. First, maximum-likelihood estimates of the η_{ij} similarity parameters in the full-choice model were computed. By the *full*-choice model, I mean the choice model with the η_{ij} parameters unconstrained except for those initial constraints stated in Equation 1. Next, a set of interpoint distances was computed from the scaling solution derived

Table 2
Best Fitting Parameters for Gaussian MDS-Choice Model

Subject		1	2	3	4	5	6	7	8
					Neutral Condition				
1	x_i	0.0000	0.6293	0.8970	1.3644	1.7026	2.0955	2.3196	2.8860
	b_i	0.7193	0.4928	0.5187	0.4765	0.4383	0.6108	0.7006	0.6895
2	x_i	0.0000	0.4032	1.0004	1.4588	2.0174	2.3012	2.7980	3.3596
	b_i	0.9005	0.9399	0.7924	0.7182	0.8152	0.7816	0.8605	0.8283
					Payoff Biased Condition				
1	x_i	0.0000	0.6931	1.1695	1.7094	2.1693	2.6852	3.0702	3.6002
	b_i	3.8728	1.8396	1.0602	0.6793	0.4127	0.2842	0.1658	0.0725
2	x_i	0.0000	0.6549	1.2557	1.8009	2.3317	2.8090	3.2495	3.9689
	b_i	7.4980	3.0156	1.3888	0.7309	0.3987	0.2721	0.1807	0.0758

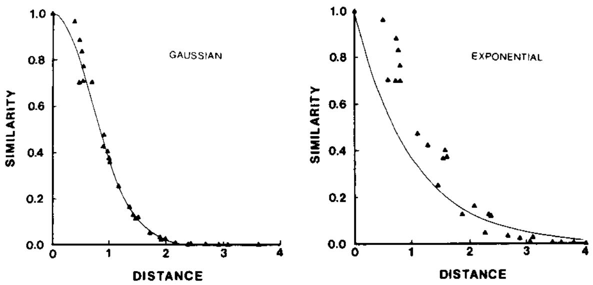

Figure 1. Graphical comparison of the Gaussian and exponential decay similarity functions on their ability to account for the similarity structure inherent in Kornbrot's data.

by fitting the MDS-choice model to the absolute identification data. Two such sets of interpoint distances were computed, one from the scaling solution in which a Gaussian similarity function was assumed and the other from the scaling solution in which an exponential decay similarity function was assumed. In the graphs shown in Figure 1, the maximum-likelihood unconstrained similarity parameters are plotted against the corresponding distance values computed from the scaling solutions. The solid curve in each graph is the theoretical similarity function. As is evident, the Gaussian similarity function captures the underlying similarity structure far better than the exponential decay similarity function. It is not surprising, therefore, that the modified Gaussian-similarity MDS-choice model yielded such improved fits to the identification data.

The reanalysis reported here raises a question on which I presently can shed little light. In addition to comparing the choice model and the Thurstone model in terms of overall goodness of fit, Kornbrot also used the method of constraint equations. The constraint equation approach suggested that the major problem with the choice model involved the *bias* constraints inherent in the general model, not the similarity constraints that were assumed in the special-case model. This suggestion is difficult to reconcile with the finding reported here, namely the dramatic improvement in overall fit achieved by using modified similarity constraints. This puzzle remains as an issue for future investigation.

Another question that arises concerns the experimental conditions that determine the best fitting similarity function. This question was studied previously by Shepard (1958b). Shepard developed an underlying process model that predicted an exponential decay similarity function under conditions of continuous reinforcement and feedback, and a Gaussian similarity function under conditions of infrequent feedback. In the experiments conducted by Kornbrot (1978) and Nosofsky (1984), however, feedback was presented on every trial. Since a Gaussian similarity function was favored in these studies, feedback cannot be the sole controlling factor. One possibility concerns whether the observed behavior is asymptotic or nonasymptotic. The experiments reported by Shepard (1958a) (in which an exponential decay function was favored) were learning studies in which subjects needed to learn the prevailing stimulus-response assignments. In contrast, Kornbrot (1978) and Nosofsky (1984) collected data from subjects who were well practiced and performing essentially at asymptote. Nosofsky (1984) fitted the MDS-choice model to several other absolute identification data sets. In all cases, the Gaussian similarity function was favored in studies that collected asymptotic performance data, whereas the exponential decay function was favored in studies that collected learning data.

Regardless of the experimental factors that determine the form of the similarity function, the main conclusion drawn from the present work is that the Gaussian MDS-choice model provides impressive fits to Kornbrot's (1978) unidimensional absolute identification data. These fits are competitive with those provided by the Thurstone model of category judgment. As stated by Kornbrot (personal communication, December 1984), "this is an interesting and intriguing result since it shows that two conceptually and mathematically different models actually provide equivalent fits to a relatively rich set of data."

REFERENCES

Kornbrot, D. E. (1978). Theoretical and empirical comparison of Luce's choice model and logistic Thurstone model of categorical judgment. *Perception & Psychophysics*, **24**, 193-208.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103-189). New York: Wiley.

Nosofsky, R. M. (1984). *Attention, similarity, and the identification-categorization relationship*. Unpublished doctoral dissertation, Harvard University.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, **22**, 325-345.

Shepard, R. N. (1958a). Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *Journal of Experimental Psychology*, **55**, 509-523.

Shepard, R. N. (1958b). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, **65**, 242-256.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, **27**, 125-140.

Thurstone, L. L. (1927). Psychophysical analysis. *American Journal of Psychology*, **38**, 368-389.

Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: Wiley.

NOTES

1. The value of x_1 may arbitrarily be set equal to 0, meaning that there are $n-1$ freely varying stimulus scale parameters. The value b_i is the bias "strength" for response R_i, with $\beta_i = b_i/\Sigma_{k=1}^{n} b_k$. Only the relative strengths are relevant, meaning that there are $n-1$ freely varying response bias parameters.

2. The empirical confusion matrices to which the models were fitted were reported by Kornbrot (1978, p. 207).

3. Following Kornbrot's procedure, the chi-square values were computed using only those cells of the judgment matrix with frequencies greater than 5.

(Manuscript received October 15, 1984;
revision accepted for publication January 18, 1985.)