

### Reason to avoid triangular designs in nonmetric scaling

MICHAEL H. BIRNBAUM  
*University of Illinois, Champaign, Illinois 61820*

Triangular designs analyzed by contemporary nonmetric scaling procedures can yield estimated scales that are nonlinearly related to the "true" scale values.

In nonmetric unidimensional scaling, observers rate the "difference," "similarity," or "dissimilarity" of stimulus pairs, and a subtractive model is used to estimate scale values that will reproduce the rank order of the judgments. For example, subjects could be asked to listen to two tones and judge the "dissimilarity" in loudness. If the subjects report that the "dissimilarity" of A to B exceeds the "dissimilarity" of C to D, then the absolute difference in scale values between A and B is assumed to exceed the difference in sensation between C and D.

#### "Differences" vs. "Dissimilarities"

There are two ways to conduct and/or analyze such an experiment. One way is to use a complete factorial design in which the first stimulus and the second stimulus are both varied and "difference" judgments are analyzed according to the equation

$$D_{ij} = J(s_j - t_i), \quad (1)$$

where  $D_{ij}$  represents the judged "difference,"  $t_i$  and  $s_j$  are the subjective values of the  $i$  and  $j$  levels of the first and second stimuli, respectively, and  $J$  is a monotonic function.

A second way to conduct the research is to present a triangular design (all possible stimulus combinations) and analyze "dissimilarity" judgments according to the equation

$$d_{ij} = M(|s_i - s_j|), \quad (2)$$

where  $d_{ij}$  is the judged "dissimilarity,"  $s_i$  and  $s_j$  are the subjective values of the stimuli, and  $M$  is a monotone function. Note that Equation 2, unlike Equation 1, assumes that  $s_i = s_j$  when  $i = j$ .

The purpose of this note is to show that the two procedures do not, in general, yield the same conclusions. In particular, if stimuli have different values depending on presentation order, then scales derived from Equation 2 can be nonlinearly related to subjective value.

Thanks are due to Robert Elmasian for suggestions on an earlier draft.

### Order/Memory Effects

Equation 2 implies that the "dissimilarity" of a stimulus to itself should be zero for all stimuli:

$$d_{ii} = 0, \text{ for all } i. \quad (3)$$

Furthermore, Equation 2 implies symmetry, that is, that presentation order has no effect:

$$d_{ij} = d_{ji}. \quad (4)$$

Equations 3 and 4 (two implications of Equation 2) are often violated when the two stimuli to be compared are presented sequentially. For example, with lifted weights, the second presentation of the same stimulus is often judged "heavier" than the memory of the first. Furthermore, memory of the first stimulus appears to depend on the average value of all stimuli presented (Parducci, Marshall, & Degner, 1966). To "avoid" these problems, researchers who use triangular designs and Equation 2 usually avoid asking subjects to compare each stimulus with itself. Usually, presentation order is counterbalanced, and  $d_{ij}$  and  $d_{ji}$  are averaged before analysis. It will be shown below that these steps do not remove the problems introduced by order effects. If the memory of the first stimulus has regressed to the mean sensation when the second stimulus is presented, then a triangular analysis will produce a distorted scale that is nonlinearly related to the "true" scale.

Suppose the first stimulus and the second stimulus have scale values that are not identical for the same stimulus, but depend on presentation order. In particular, suppose that the subjective value of the first stimulus is linearly related to the subjective value of the same stimulus when presented second:

$$t_i = as_i + b, \quad (5)$$

where  $s_i$  and  $t_i$  are the scale values of stimulus  $i$  when presented second or first, respectively, and  $a$  and  $b$  are constants. The linear function, with  $a < 1$ , might represent a regression of the scale value of the first stimulus in memory to the average sensation.

#### Theory of "Dissimilarity"

When subjects are asked to compare two stimuli (e.g., lift two weights successively) and judge the "dissimilarity" in heaviness, it seems reasonable to suppose that they perceive the algebraic difference (i.e., they know which stimulus seems heavier) and that they convert the difference to an absolute value when judging "dissimilarity." Thus,

$$d_{ij} = M|J^{-1}(D_{ij})|, \quad (6)$$

where  $M$  is a function. Note that  $M$  need not have an inverse function (two algebraic differences of opposite sign have the same absolute difference).

**Self-Comparisons**

Substituting Equation 5 into Equation 1 yields

$$D_{ij} = J(s_j - as_i - b). \tag{7}$$

Equation 7 implies that the judged "difference" between a stimulus and itself will be given by the equation

$$D_{ii} = J[s_i(1 - a) - b], \tag{8}$$

which implies that, unless  $a = 1$ , Equation 3 (and therefore Equation 2) will be violated. In particular, "dissimilarity" of a stimulus to itself should depend on the absolute deviation of a stimulus from the central value.

**Adjacent Comparisons**

Equation 7 implies that the judged "difference" between adjacent stimulus levels will be given by the formulas:

$$D_{ik} = J(s_{i+1} - as_i - b) \tag{9}$$

and

$$D_{ki} = J(s_i - as_{i+1} - b), \tag{10}$$

where  $k = i + 1$ . Let the differences between successive scale values be denoted  $\Delta_i = s_{i+1} - s_i$ . Then Equations 9 and 10 become

$$D_{ik} = J[\Delta_i + (1 - a)s_i - b] \tag{11}$$

and

$$D_{ki} = J[-a\Delta_i + (1 - a)s_i - b]. \tag{12}$$

If there were no presentation order effect, then  $a = 1$  and Equations 11 and 12 would depend only on  $\Delta_i$ , as in Equation 2. However, Equations 11 and 12 show that if  $a < 1$ , then judged "differences" between adjacent stimuli will be biased *upward* in proportion to the subjective value of  $s_i$ . Note that the bias will be in the same direction irrespective of stimulus presentation order.

Analysis by Equation 2 based on a triangular design will therefore yield scale values that are non-linearly related to those that would be derived from the entire factorial design. In particular, if  $a$  is less than 1 and if the stimuli are actually equally spaced in subjective value, the usual triangular analysis (Equation 2) will yield scale value estimates that are positively accelerated relative to the values derived from Equation 1. When  $a$  is greater than 1, the non-linear distortion will be in the opposite direction.

**Numerical Example**

Table 1 was generated from Equation 7, using the successive even integers from 0 to 16 as scale values,  $a = .5$ ,  $b = 4$ , and  $J$  was an identity function. Marginal means for the row (A) and column (B) values, which represent estimates of scale values according to Equation 2 (if  $J$  is linear), are linearly related to each other. Furthermore, nonmetric analysis in terms of Equation 1 yields scale values that are linearly related to each other (and to the marginal means). In other words, analysis by means of Equation 1 correctly yields scale values that are equally spaced.

In the example, all of the  $\Delta_i$  are equal because the scale values are equally spaced. However, above the diagonal in Table 1, the judged "dissimilarity" of adjacent stimulus levels increases as one moves up the scale. Thus,  $d_{3,4} < d_{4,5} < d_{5,6} < \dots < d_{8,9}$ . In the ordinary triangular analysis (Equation 2), this result would be interpreted to mean that successive differences between scale values increase as one moves up the scale. In this case, that interpretation would be incorrect. This result is predicted by Equation 11.

**Table 1**  
Hypothetical Data Generated From the Equation  $D_{ij} = s_j - .5s_i$

First Stimulus A	Second Stimulus B									A*
	1	2	3	4	5	6	7	8	9	
1	-4	-2	0	2	4	6	8	10	12	4
2	-5	-3	-1	1	3	5	7	9	11	3
3	-6	-4	-2	0	2	4	6	8	10	2
4	-7	-5	-3	-1	1	3	5	7	9	1
5	-8	-6	-4	-2	0	2	4	6	8	0
6	-9	-7	-5	-3	-1	1	3	5	7	-1
7	-10	-8	-6	-4	-2	0	2	4	6	-2
8	-11	-9	-7	-5	-3	-1	1	3	5	-3
9	-12	-10	-8	-6	-4	-2	0	2	4	-4
B*	-8	-6	-4	-2	0	2	4	6	8	

Note—It is assumed that  $s_i = s_j$  when  $i = j$ . Marginal means for B (column means) are scale values. Marginal means for A in this case are .5s. \*Marginal mean.

When Equation 2 was fit to the upper triangular design in Table 1 (where  $B - A > 0$ ), the estimated scale values were  $-.84, -.65, -.52, -.36, -.18, .08, .40, .81$ , and  $1.27$ , which are a positively accelerated function of the actual scale values (from 0 to 16, respectively) used to generate the data. (The computer program, TORSCA, was used for the analysis. Since the entries in Table 1 represent algebraic rather than absolute differences, the absolute values of  $D_{ij}$  were used for the TORSCA runs.) The nonlinearity between the actual scale values and derived values is predicted from Equations 6 and 11.

When TORSCA was applied to the lower triangle of Table 1 (where  $B - A < 0$ ), the estimated scale values were  $-1.27, -.81, -.40, -.08, .18, .36, .52, .65$ , and  $.84$ , respectively. Thus, the scale values that would be derived from the data in Table 1 depend on whether the upper or lower triangle was used. Similar results were obtained when the value of  $a$  was set to the less extreme value of  $.9$ .

Table 1 also shows that  $D_{ii}$  are not all zero, but that small stimuli seem "smaller" and large stimuli seem "larger" than their respective memories. Of course, "dissimilarity" judgments would show that extreme stimuli are most "dissimilar" to themselves, contrary to Equation 3, but as predicted by Equation 8.

### Discussion

It might be thought that if stimulus order and position are counterbalanced, then this difficulty would be overcome. Indeed, averaging  $|D_{ij}|$  and  $|D_{ji}|$  in Table 1 would yield a triangular matrix from which a scale would be derived that is linearly related to the original scale. (The first two diagonals, where  $|i - j| = 0$  and  $|i - j| = 1$ , must be excluded!) However, if  $J$  is nonlinear, then averaging  $d_{ij}$  and  $d_{ji}$  would not generally cancel the distortion. Because nonmetric scaling allows nonlinear  $J$  (or  $M$ ) functions, it should not be assumed that counterbalancing and averaging removes the distortion.

Unfortunately, those who have used triangular designs have not typically asked subjects to compare each stimulus with itself. These judgments (Equation 8; major diagonal in Table 1) can provide evidence of a regression effect that vitiates the use of Equation 2. Thus, scales and conclusions derived in that research remain suspect.

In previous research with factorial analyses of "difference" judgments, it has been possible to derive separate scales for each order (or position) (Birnbbaum, 1978, 1980). It has been found that when stimuli are presented simultaneously (e.g., subject lifts two weights simultaneously, one in each hand, or com-

pares two simultaneously available visual stimuli), scale values for the same stimulus are nearly the same for both positions. However, when stimuli are presented sequentially (e.g., two tones presented one after the other), scale values for the same stimulus depend on whether the stimulus was first or second. (Birnbbaum & Elmasian, 1977; Elmasian & Birnbbaum, Note 1). Therefore, the problem raised in this paper seems a real threat for experiments in which the two stimuli to be compared are presented sequentially.

It should be mentioned that if the subjective value of the second stimulus depends on the values of both the first and second stimuli in a nonlinear way, then even the factorial analysis (as described above) will have difficulty. Such effects have occurred with auditory stimuli presented with short interstimulus intervals (e.g., Elmasian, Galambos, & Bernheim, 1980). Such interactions would be expected to produce deviations from the subtractive model when the interstimulus interval is less than half a second. With 1-sec interstimulus intervals, Birnbbaum and Elmasian (1977) found that any nonlinear effects were small enough to be neglected.

In sum, nonmetric scaling based on triangular designs may yield psychophysical scales that are nonlinearly related to "actual" psychophysical scales. This issue is also relevant to "ratio" judgments and to multidimensional scaling of "similarity" judgments. The problem can be remedied by use of factorial designs and analyses that permit stimuli in different orders or positions to have different scale values in the analysis.

### REFERENCE NOTE

1. Elmasian, R., & Birnbbaum, M. H. *A harmonious note on pitch*. Unpublished manuscript, 1979.

### REFERENCES

- BIRNBAUM, M. H. Differences and ratios in psychological measurement. In N. J. Castellan, Jr., & F. Restle (Eds.), *Cognitive theory* (Vol. 3). Hillsdale, N.J.: Erlbaum, 1978.
- BIRNBAUM, M. H. Comparison of two theories of "ratio" and "difference" judgments. *Journal of Experimental Psychology: General*, 1980, **109**, 304-319.
- BIRNBAUM, M. H., & ELMASIAN, R. Loudness "ratios" and "differences" involve the same psychophysical operation. *Perception & Psychophysics*, 1977, **22**, 383-391.
- ELMASIAN, R., GALAMBOS, R., & BERNHEIM, A. Loudness enhancement and decrement in four paradigms. *Journal of the Acoustical Society of America*, 1980, **67**, 601-607.
- PARDUCCI, A., MARSHALL, L. M., & DEGNER, M. Interference with memory for lifted weight. *Perception & Psychophysics*, 1966, **1**, 83-86.

(Received for publication February 12, 1981;  
accepted February 12, 1981.)