A REPEATED-MEASURES PERMUTATION PROCEDURE FOR ANALYZING SEMANTIC DIFFERENTIAL DATA

KENNETH J. BERRY AND PAUL W. MIELKE, JR.

Colorado State University

Request reprints from K. J. Berry, Department of Sociology, Colorado State University, Fort Collins, CO 80523-1784

A REPEATED-MEASURES PERMUTATION PROCEDURE FOR ANALYZING SEMANTIC DIFFERENTIAL DATA

Summary.—Data from the semantic differential are often analyzed with a repeated-measures analysis of variance. The problem of an appropriate distance function is addressed and a permutation multivariate repeated-measures analysis of variance is introduced which preserves the Euclidean data space mandated by the semantic differential.

A REPEATED-MEASURES PERMUTATION PROCEDURE FOR ANALYZING SEMANTIC DIFFERENTIAL DATA

Since its introduction in 1952 (Osgood, 1952; Osgood & Suci, 1952) and extensive elaboration in 1957 (Osgood, Suci, & Tannenbaum, 1957), the semantic differential technique for investigating connotative and metaphorical meanings has played an important role in psychological research (Heise, 1969; Finstuen, 1977). Over the past four decades the semantic differential has been applied in many other areas of research including education, marketing, communication research, and cross-cultural studies. It is rarely mentioned in recent discussions of the semantic differential that the method radically departs from the original formulation as proposed by Osgood, *et al.* (1957). The purpose of this report is to propose an alternative method of analysis for the semantic differential which conforms to the procedure initially advanced by Osgood, *et al.* (1957).

The semantic differential is based on subjects' ratings of a number of concepts (e.g., Mother, Fate, Success) on a set of bipolar adjective scales (e.g., severe-lenient, active-passive, rational-intuitive). Osgood, *et al.* (1957) factor analyzed the scales into the three well-known dimensions of Evaluation, Potency, and Activity. These three dimensions are the ones usually employed in semantic differential research and have been replicated and corroborated in many different studies and populations. The rationale underlying the application of the semantic differential to an affective domain requires that the semantic space be conceived of as a Euclidean region of unknown dimensionality as defined by the set of bipolar adjective scales (Osgood, *et al.*, 1957, p. 25). Because it is possible to ascertain subject differences only with respect to specified dimensions (Cronbach & Gleser, 1953), the degree of difference among subjects' score sets is calculated by the generalized distance formula in an r-dimensional geometry where each score set is represented by a point in the r-dimensional semantic space. The differences between subject score sets i and j in the r-dimensional space are indexed by

$$D_{ij} = \left[\sum_{k=1}^{r} (x_{ki} - x_{kj})^2 \right]^{\frac{1}{2}} \qquad [1]$$

(Osgood, *et al.*, 1957, p. 91). Thus, the difference between a pair of subjects is proportional to the Euclidean distance between them, with high scores indicating large differences. Although D^2_{ij} can be used as the measure of difference, it is preferable to utilize the stipulated Euclidean distance measure, i.e., D_{ij}, because (a) D_{ij} reflects the natural Euclidean space of the data, i.e., the data space; (b) larger differences between subject score sets are greatly exaggerated with D_{ij}^2; (c) the property of additivity of distances, a necessary condition of a ratio scale, is destroyed (Cliff, 1959); (d) the triangle inequality (Bock, 1975, p. 29) is not satisfied; and (e) the data space is converted to a nonmetric analysis space (Mielke, 1986).

Because D_{ij} is not normally distributed (Cronbach & Gleser, 1953, p. 459; Osgood, *et al.*, 1957, p. 101) and because a Euclidean distance measure is mandated (Osgood & Suci, 1952; Osgood, *et al.*, 1957; Brown, 1958; Carroll, 1959), a test which does not assume normality and which is capable of mapping a Euclidean semantic space with an r-dimensional Euclidean distance measure is ideally suited to analyzing results of semantic differential studies.

METHOD

The semantic differential technique leads naturally to analysis of variance with repeated-measures (Bynner & Coxhead, 1979) and repeated-measures experimental designs are frequently employed in semantic differential research (e.g., DeBurger & Donahoe, 1965; Johnson, Thomson, & Frinke, 1960; Osgood & Luria, 1954). The immediate impact of analyzing semantic differential scores with any analysis of variance is to replace the requisite Euclidean distances (D_{ij}) with squared distances (D_{ij}^2). This introduces an analysis space based on squared distances which is totally different from the previously defined and desired Euclidean data space. It is absolutely essential that the analysis space and the data space correspond exactly and that the congruence principle be satisfied (Mielke & Berry, 1983). The required r-dimensional semantic-differential data space can be preserved and analyzed with a permutation test based on Euclidean distances between semantic differential scores.

Consider a one-factor repeated-measures experimental design with n subjects, g treatments, and r observations on each subject. Let x_{kij} denote the semantic differential score of the k^{th} response ($k = 1,\ldots,r$) on the i^{th} treatment ($i = 1,\ldots,g$) for the j^{th} subject ($j = 1,\ldots,n$). An inherent component of the variability in all repeated-measures experimental designs is that due to between-subject differences. In any analysis, this between-subject variability must be controlled so that the treatment effects may be accurately assessed. Alignment of the individual semantic differential variables for each subject is a technique which controls for the between-subject variability (Mielke & Iyer, 1982). Alignment is accomplished by replacing x_{kij} with $x_{kij} - \tilde{x}_{kj}$ where $\tilde{x}_{kj}$ is the median of the g treatment responses $(x_{k1j},\ldots,x_{kgj})$ for $k = 1,\ldots,r$ and $j = 1,\ldots,n$. The adjusted semantic differential measurements are thus aligned to zero for all subjects. Let $x_{ij}' = (x_{1ij}, \ldots, x_{rij})$ denote a vector of r aligned observations associated with the i^{th} treatment and the j^{th} subject. For $v > 0$, let

$$\Delta(x,y) = \left[\sum_{k=1}^{r}(x_k - y_k)^2\right]^{\frac{v}{2}} \qquad [2]$$

designate a measure of difference, i.e., distance, between the r-dimensional points $x' = (x_1, \ldots, x_r)$ and $y' = (y_1, \ldots, y_r)$. Note that when $v = 1$, $\Delta(x,y)$ of Equation 2 is the ordinary Euclidean distance, D_{ij}, of Osgood, *et al.* (1957) given in Equation 1. The statistic of interest is then given by

$$\delta = \left[g\binom{n}{2}\right]^{-1} \sum_{i=1}^{g} \sum_{j<k} \Delta(x_{ij}, x_{ik}) \qquad [3]$$

where $\sum_{j<k}$ is the sum over all j and k such that $1 \le j < k \le n$. Since the n subjects are specified, the randomization associated with a repeated-measures experiment is confined to all permutations of the g treatments on each of the n subjects. The null hypothesis (H_0) states that the distribution of δ assigns an equal probability to each of the

$$M = (g!)^n \qquad [4]$$

possible allocations of the g r-dimensional response measurements to the g treatment positions for each of the n subjects. A permutation test determines the exact probability of a realized test statistic by computing the proportion of all possible test statistic values equal to or greater than the realized value (Fisher, 1935; Hoeffding, 1952; Hubert & Levin, 1977; Pitman, 1937, 1938; Wald & Wolfowitz, 1944; Welch, 1937). Thus, the probability (P) associated with an observed value of δ (δ_o) is the probability under H_0 of observing a value of δ equal to or less than δ_o. The exact probability of an observed δ is given by

$$P(\delta \le \delta_o | H_0) = \frac{\text{number of } \delta\text{'s} \le \delta_o}{M} . \qquad [5]$$

Although a permutation test provides an exact probability value, each application of the test requires the complete enumeration of all possible permutations of the realized data set. Consequently, notwithstanding the use of modern high-speed computers, imaginative conceptualization, and inventive programming (Chung & Fraser, 1958; Edgington, 1980; Green, 1977), computation is impractical except for very small samples, e.g., a one-factor repeated-measures design with $g = 3$ treatments on $n = 15$ subjects generates $M = 4.7018 \times 10^{11}$ permutations. It is for this reason that approximate randomization tests, based on a random sample of all possible permutations, have become popular. There are two disadvantages to these approximate randomization tests. First, there will necessarily be an increase in the amount of type I error, due to the additional sampling. Second, and most important, is the issue of replication. Two researchers using the same data set and identical protocols will necessarily arrive at different results, due to the different samples drawn for analysis. Indeed, there is nothing to prohibit the unscrupulous investigator from drawing repeated randomizations until one is found that will support the research hypothesis. An alternative method which avoids the excessive computational demands of an exact test and the additional type I error of an approximate randomization test, is a moment aproximation procedure which approximates the discrete probability distribution of an exact test with a continuous distribution that fits the lower exact moments of the discrete probability distribution.

If δ_j denotes the j^{th} value among the M possible values of δ, then the mean, variance, and skewness of δ, under H_0, are given by

$$\mu_\delta = \frac{1}{M}\sum_{j=1}^{M} \delta_j, \qquad [6]$$

$$\sigma_\delta^2 = \frac{1}{M}\sum_{j=1}^{M}(\delta_j - \mu_\delta)^2, \qquad [7]$$

and

$$\gamma_\delta = \left[\frac{1}{M}\sum_{j=1}^{M}(\delta_j - \mu_\delta)^3\right] \Big/ \sigma_\delta^3, \qquad [8]$$

respectively. Efficient computational techniques for obtaining μ_δ, σ_δ^2, and γ_δ are described in detail by Mielke and Iyer (1982). The standardized test statistic for an observed δ, δ_o, is given by

$$T_o = \frac{\delta_o - \mu_\delta}{\sigma_\delta} \qquad [9]$$

and is approximated, under H_0, by the Pearson type III distribution with a density function given by

$$f(y) = \frac{(-2/\gamma_\delta)^{4/\gamma_\delta^2}}{\Gamma(4/\gamma_\delta^2)} \{-(2 + y\gamma_\delta)/\gamma_\delta\}^{(4-\gamma_\delta^2)/\gamma_\delta^2} e^{-2(2+y\gamma_\delta)/\gamma_\delta^2} \qquad [10]$$

where $-\infty < y < -2/\gamma_\delta$. The probability value for δ_o is given by

$$P(\delta \le \delta_o | H_0) \doteq \int_{-\infty}^{T_o} f(y)\, dy. \qquad [11]$$

EXAMPLE

Consider a study which employs the semantic differential as an operational index of change in the representational mediation process during therapy (cf, Endler, 1961). Specifically, the research seeks to analyze changes in the meaning of the clients' "me" concept during therapy and concentrates on the Evaluative, Potency, and Activity factors of meaning. Fifteen adult subjects who were in psychotherapy for personal or emotional problems were administered the Osgood Semantic Differential before beginning therapy, after six months of therapy, and immediately upon completion of therapy. To test the H_0 of no change in the "me" concept during therapy, i.e., the therapy had no effect on the subjects, the data is analyzed with a Euclidean analysis. The raw data are presented in Table 1.

Insert Table 1 about here

Analysis of the representative data set in Table 1 with $v = 1$ yields $\delta = 0.5237$, $\mu_\delta = 0.6038$, $\sigma_\delta^2 = 1.2467\times10^{-4}$, $\gamma_\delta = -0.9658$, $T_o = -7.1659$, and $P = 1.1070\times10^{-5}$. The analysis indicates that the meaning of "me" (as measured by the evaluative, Potency, and Activity factors of the semantic differential) is significantly modified during therapy.

Iyer, Berry, and Mielke (1983) provide a FORTRAN program to compute all the results obtained in this paper. Copies of the program are available from K. J. Berry, Department of Sociology, Colorado State University, Fort Collins, CO 80523-1784, from P. W. Mielke, Jr., Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877, or by e-mail from Internet addresses: berry@lamar.colostate.edu or mielke@lamar.colostate.edu.

TABLE 1

SEMANTIC DIFFERENTIAL SCORES ON THE EVALUATIVE (E),

POTENCY (P), AND ACTIVITY (A) DIMENSIONS

Subject	Treatment 1			Treatment 2			Treatment 3		
	E	P	A	E	P	A	E	P	A
1	4.3	5.5	3.9	4.6	5.4	4.1	4.1	5.6	4.2
2	5.9	4.5	2.7	6.0	4.3	2.8	5.6	4.6	3.0
3	3.0	6.1	6.2	3.1	6.0	6.3	2.9	6.3	6.4
4	5.4	6.7	4.7	5.6	6.6	4.8	5.3	6.9	4.9
5	4.1	2.8	5.3	3.5	4.2	4.0	3.9	2.9	5.6
6	2.9	4.3	6.1	3.2	4.1	6.4	2.7	4.4	6.5
7	3.6	5.6	4.8	3.8	5.5	4.9	4.1	4.4	2.7
8	3.3	4.0	5.1	3.4	3.9	5.3	3.2	4.2	5.4
9	2.7	6.1	3.5	2.8	5.8	3.6	2.6	6.2	3.8
10	4.3	2.9	5.0	4.5	2.7	5.1	4.1	3.2	5.2
11	4.4	5.3	4.4	4.5	5.2	4.6	4.3	5.4	4.8
12	3.9	4.8	5.1	2.4	5.5	4.5	3.8	4.9	5.4
13	5.0	3.1	6.2	5.1	3.0	6.4	6.5	2.4	4.5
14	6.2	4.8	6.1	6.3	4.5	6.2	6.0	4.9	6.3
15	5.3	4.9	3.7	5.5	4.7	3.8	5.2	5.1	4.1

REFERENCES

BOCK, R. D. (1975) *Multivariate statistical methods in behavioral research*. New York, NY: McGraw-Hill.

BROWN, R. W. (1958) [Review of the book *The measurement of meaning*]. *Contemporary Psychology*, 3, 113-115.

BYNNER, J., & COXHEAD, P. (1979) Some problems in the analysis of semantic differential data. *Human Relations*, 32, 367-385.

CARROLL, J. B. (1959) [Review of the book *The measurement of meaning*]. *Language*, 35, 58-77.

CHUNG, J. H., & FRASER, D. A. S. (1958) Randomization tests for a multivariate two-sample problem. *Journal of the American Statistical Association*, 53, 729-735.

CLIFF, N. (1959) Adverbs as multipliers. *Psychological Review*, 66, 27-44.

CRONBACH, L. J., & GLESER, G. C. (1953) Assessing similarity between profiles. *Psychological Bulletin*, 50, 456-473.

DeBURGER, R. A., & DONAHOE, J. W. (1965) Relationship between the meanings of verbal stimuli and their associative responses. *Journal of Verbal Learning and Verbal Behavior*, 4, 25-31.

EDGINGTON, E. S. (1980) *Randomization tests*. New York, NY: Marcel Dekker.

ENDLER, N. S. (1961) Changes in meaning during psychotherapy as measured by the semantic differential. *Journal of Counseling Psychology*, 8, 105-111.

FINSTUEN, K. (1977) Use of Osgood's semantic differential. *Psychological Reports*, 41, 1219-1222.

FISHER, R. A. (1935) *The design of experiments*. Edinburgh: Oliver & Boyd.

GREEN, B. F. (1977) A practical interactive program for randomization tests of location. *American Statistician*, 31, 37-39.

HEISE, D. R. (1969) Some methodological issues in semantic differential research. *Psychological Bulletin*, 72, 406-422.

HOEFFDING, W. (1952) The large-sample power of tests based on permutations of observations. *Annals of Mathematical Statistics*, 23, 169-192.

HUBERT, L. J., & LEVIN, J. R. (1977) Inference models for categorical clustering. *Psychological Bulletin*, 83, 878-887.

IYER, H. K., BERRY, K. J., & MIELKE, P. W. (1983) Computation of finite population parameters and approximate probability values for multi-response randomized block permutation procedures (MRBP). *Communications in Statistics—Simulation and Computation*, 12, 479-499.

JOHNSON, R. C., THOMSON, C. W., & FRINCKE, G. (1960) Word values, word frequency, and duration thresholds. *Psychological Review*, 67, 332-342.

MIELKE, P. W. (1986) Non-metric statistical analysis: some metric alternatives. *Journal of Statistical Planning and Inference*, 16, 430.

MIELKE, P. W., & BERRY, K. J. (1983) Asymptotic clarifications, generalizations, and concerns regarding an extended class of matched pairs tests based on powers of ranks. *Psychometrika*, 48, 483-485.

MIELKE, P. W., & IYER, H. K. (1982) Permutation techniques for analyzing multi-response data from randomized block experiments. *Communications in Statistics - Theory and Methods*, A11, 1427-1437.

OSGOOD, C. E. (1952) The nature and measurement of meaning. *Psychological Bulletin*, 49, 197-237.

OSGOOD, C. E., & LURIA, Z. (1954) A blind analysis of a case of triple personality using the semantic differential. *Abnormal and Social Psychology*, 49, 579-591.

OSGOOD, C. E., & Suci, G. J. (1952) A measure of relation determined by both mean difference and profile information. *Psychological Bulletin*, 49, 251-262.

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. (1957) *The measurement of meaning*. Urbana, IL: University of Illinois Press.

PITMAN, E. J. G. (1937) Signficance tests which may be applied to samples from any populations. *Supplement of the Journal of the Royal Statistical Society*, B4, 119-130.

PITMAN, E. J. G. (1938) Significance tests which may be applied to samples from any populations: III. The analysis of variance test. *Biometrika*, 29, 322-335.

WALD, A., & WOLFOWITZ, J. (1944) Statistical tests based on permutations of the observations. *Annals of Mathematical Statistics*, 15, 358-372.

WELCH, B. L. (1937) On the z-test in randomized blocks and Latin squares. *Biometrika*, 29, 21-52.