A VARIABILITY METHOD FOR OBTAINING SUBJECTIVE INFORMATION

DAVID R. RINK, Indiana University at Kokomo
JAMES E. DUNN, University of Arkansas

1993 Richard D. Irwin Distinguished Paper Award

One efficient data collection methodology for obtaining subjective information has received little attention. It is balanced incomplete block (BIB) designs. The authors demonstrate how to develop and apply a BIB design.

INTRODUCTION

Numerous stimulus-centered methods exist for collecting judgmental data (Green and Tull 1975). But, only paired comparison and rank order techniques have been extensively used (Wiley 1978). However, major methodological questions can be raised regarding both approaches. The method of paired comparisons necessitates a large number of observations, and deriving all pairs of stimuli "becomes very tedious and time-consuming when the number of stimuli becomes large." Not only does the number of pairs increase much faster than the number of stimuli, but so does "the amount of overdetermination, ... and, hence the number of degrees of freedom left after fitting the data" (Torgerson 1958, p. 191). Furthermore, with a redundant data collection method as in various multidimensional scaling and conjoint measurement models, the respondent is likely to become tired and bored (Green and Carmone 1970; Green 1974), which can affect adversely the reliability and accuracy of the study's results and conclusions. In the complete rank order approach, the respondent may become tired if he/she is asked to rank a large number of items.

GENERAL NATURE OF BIB DESIGNS

The problems of respondent fatigue and boredom can be alleviated by using one of the balanced incomplete block (BIB) designs (Cochran and Cox 1957). When the number of stimuli is large and the sample size is limited, it may be impractical or impossible for all stimuli to be applied to each respondent, and an incomplete block design may be used. If the design is balanced so that every respondent (or block) contains k experimental units, every stimuli (or treatment) appears in r blocks, and every stimuli (or treatment) appears with every other treatment an equal number of times, the design is then called a balanced incomplete block (BIB) design (Conover 1971, pp. 275-276). These designs are suitable in cases involving "subjective ranking by a small panel of judges for the detection of differences" (Bradley and Terry 1952, p. 335); that is, in situations where "individuals are asked to make a comparative rating of different objects that are presented to them" (Cochran and Cox 1957, p. 440). BIB designs "still permit various consistency checks to be carried out but reduce considerably the number of independent pairwise comparisons which the subject is forced to make" (Green and Carmone 1970, p. 85); and they "provide a type of balancing across stimulus pairs and subjects" (Green and Tull 1975, p. 178).

The Journal of Marketing Management
Volume 3, Issue 2, pages 1-5

Every BIB design must satisfy *both* of these conditions (Federer 1955, p. 416):

$$tr = kb \qquad (1)$$

$$(t-1)\,\lambda = r(k-1) \qquad (2)$$

where t = number of treatments to be examined
k = number of experimental units per block ($k < t$)
b = total number of blocks
r = number of times each treatment appears ($r < b$)
λ = the number of blocks in which the i^{th} treatment and the j^{th} treatment appear together (λ is the same for all pairs of treatments).

ANALYTICAL PROCEDURE FOR BIB DESIGNS

The traditional technique used for analyzing data in BIB designs is the analysis of variance. But, "since all treatments do not appear in the same block of BIB designs, treatment effects are partially confounded with block effects and vice versa." Therefore, a roundabout procedure must be used to "obtain estimates of treatment effects free of block effects." This specific procedure and the analyses of variance for various BIB designs may be found in several sources (Banks 1974; Cochran and Cox 1957).

An analytical procedure that is computationally simpler incorporates the Durbin Test (Durbin 1951), Coefficient of Concordance (Gibbons 1971, pp. 250-262; Kendall 1955, pp. 94-116), and Guttman scale (Guttman 1946). Because the observations in BIB designs may consist merely of ranks and therefore would not meet the normality assumptions required for applying parametric techniques, the Durbin test is appropriate. These two assumptions underlie the use of the Durbin test in BIB designs: blocks (or judges) are mutually independent of each other; and within each block, the observations may be arranged in increasing order according to some criterion of interest (Conover 1971, p. 277). The null hypothesis commensurate with the Durbin test states: Each ranking of the random variables within each block is equally likely. That is, the treatments have identical effects. Alternatively, at least one treatment tends to yield larger observed values than at least one other treatment (Conover 1971, p. 277).

The Durbin test statistic used in BIB designs is defined as (Conover 1971, p. 277):

$$T=\frac{12(t-1)}{rt(k-1)(k+1)}\sum_{j=1}^{t} R_j^2-3\left[\frac{r(t-1)(K+1)}{k-1}\right] \qquad (3)$$

where $R_j=\sum_{i=1}^{b} R(X_{ij})$ = sum of ranks assigned to the r observed values under the jth treatment.

In terms of a decision rule, the null hypothesis should be rejected at the alpha level of significance if the Durbin test statistic T exceeds the (1- α)th quantile of a chi-square random variable with t-1 degrees of freedom (Conover 1971, p. 277).

Assuming the rejection of the null hypothesis, a Guttman scale can be derived. The researcher can then determine those objects (or treatments) respondents (or blocks) considered most important as well as ascertain the intensity with which respondents rank ordered objects. The formula used to measure agreement among observers regarding their rankings of objects in BIB designs is defined as (Gibbons 1971, p. 260; Kendall 1955, p. 105):

$$W=\frac{12\sum_{j=1}^{t} R_j^2-3r^2t(k+1)^2}{\lambda^2t(t^2-1)} \qquad (4)$$

A MARKETING APPLICATION

The president of a multinational corporation wanted his top managers to rank 15 potential warehouse location sites from most preferred to least preferred. (The corporation could fund only three of these sites.) Because these executives would probably find it difficult to rank this many sites, the president decided to adopt a BIB design. The president felt the executives could effectively rank seven location sites. Consulting a table of BIB designs (Raghavaro 1971), the president selected Series 43 as the optimal BIB design to adopt. Operationally, Series 43 means (Table 1): 15 sites (treatments) will be evaluated (t=15); 15 executives (blocks) will be sampled (b=15); each location site will be repeated seven times (r=7) in the BIB design; each executive will rank seven sites (k=7); and each site will be compared with every other site by three executives (λ=3).

After randomizing the order of location sites presented to each executive and the assignment of executives to blocks, the president conducted his survey. The results are summarized in Table 1. After summing the ranks for each location site, the Durbin test statistic was computed using formula (3):

$$T=\frac{12(5-1)}{7(15)(7-1)(7+1)}[(39)^2+(35)^2+\ldots+(19)^2]$$

$$-3\left[\frac{7(15-1)(7+1)}{(7-1)}\right]$$

$$=25.5$$

For t-1 or 14 degrees of freedom, this Durbin T value was statistically significant at the .0299 level. Therefore, given an alpha of .05, the null hypothesis that treatments (or location sites) were identically preferred was rejected. At least one site tended to yield larger observed values than at least one other site.

Next, the Coefficient of Concordance was calculated using formula (4):

$$W = \frac{12[(39)^2 + (35)^2 + \ldots + (19)^2] - 3(7)^2(15(7+1)^2}{(3)^2(15)[(15)^2 - 1]}$$

$$= .30357$$

This value of W was statistically significant at the same level as the Durbin test statistic, because the Coefficient of Concordance is a linear transformation of the Durbin test statistic. The result of a significant W was interpreted to mean that the rankings of 15 location sites across the 15 executives exhibited some degree of consistency. Furthermore, these results indicated the executives were using the same criterion in evaluating sites.

Finally, to more clearly ascertain the exact order and spread among the location sites, thereby indicating the intensity with which executives ranked sites, a Guttman scale was derived. The methodology for deriving the Guttman scale for a set of ordinal data is presented by Guttman (1946), and will not be discussed here. The last row of Table 1 summarizes the Guttman scale scores for the 15 location sites. If the most preferred site received a rank of "1" while the least preferred site was awarded a rank of "7" by each executive, then those sites possessing the most negative Guttman scale scores would represent the most preferred warehouse location sites across *all* respondents. In this situation, the president concluded: Location site 10 (-.49) was overwhelmingly preferred to the other 14 sites by his top managers; although somewhat removed from site 10, sites 15 (-.28) and 12 (-.28) were equally preferred; and site 9 (-.23) was somewhat removed from sites 15 and 12 in terms of preference. Hence, the president decided to fund warehouse location sites 10, 15, and 12 (in order of preference).

ADVANTAGES AND LIMITATIONS OF BIB DESIGNS

Several advantages are associated with the use of BIB designs. First, although the number of comparisons per respondent are reduced, the phenomenon of "balancing" across subjects enables them as a group to rank many objects (Green and Tull 1975, p. 178). Second, the researcher is no longer bound unequivocally to the restriction that the number of treatments must equal the number of experimental units in a block (Yates 1936,

TABLE 1
RESULTS FROM BIB DESIGN FOR WAREHOUSE LOCATION SITE PREFERENCE STUDY

Exec #	Warehouse Location Site Number														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	4									2	6	3	5	7	1
2		7					1	4	6				2	5	3
3			2			6		7	3		4	1			5
4				2		1	4		7	6		5		3	
5					3	5	2	4		1	6		7		
6			1	3	5	7							4	6	2
7		1		6	2		3				7	5			4
8		1	4		7			6		3		2		5	
9		7	4	1					2	3	5		6		
10	7			5	6			4	3	1					2
11	6		1		3		7		4		5			2	
12	5		4	3			2	6				1	7		
13	6	5			4	1			2			3	7		
14	5	7		1		2		6			3			4	
15	6	7	5			3	4			1					2
R_j's	39	35	21	21	30	25	23	37	27	17	36	20	38	32	19
Guttman	.33	.36	-.02	-.19	.12	-.04	-.15	.19	-.23	-.49	.10	-.28	.41	.18	-.28

p. 122), which is the major characteristic of randomized block and Latin square designs. BIB designs allow the number of treatments to exceed the number of experimental units per block. Third, the procedure by which the rankings of several judges are combined "permits an overall test of significance without the usual assumption that members of a panel agree upon the nature of the differences to be detected" (Bradley and Terry 1952, p. 325). Fourth, the use of BIB designs increases the precision of the study, even with a small sample size. This is because "blocking" and replication reduce variability. That is, if the population tends to be homogeneous in terms of some variable of interest, then a small sample interviewed in detail will probably result in more valid inferences. On the other hand, with a heterogeneous population, there is a possibility of too much variation in subjects' rankings. From an administration standpoint, BIB designs save the respondent time, which may make possible a higher completion rate as well as cut interview costs. Sixth, "distribution-free" or nonparametric statistical techniques may be used in BIB designs. These techniques are easier to learn and use than their parametric counterparts. Furthermore, the nonparametric procedures are appropriate for most marketing problems where the assumption of normality is unwarranted. Finally, as long as the observations (even if non-numeric) are amenable to rank-ordering according to some criterion of interest (Conover 1971, p. 276), the researcher does not have to invent a scale of measurement he/she can use. This is the *major* advantage of the rank-order procedure.

The major limitation of BIB designs is when the number of treatments is large, the number of replications required also tends to be large (Yates 1936). Second, unlike the method of paired comparisons, once the number of blocks (or respondents) is ascertained in BIB designs, the researcher *must* acquire this exact number—no more, no less. Finally, BIB designs are applicable only to situations in which respondents' preferences are homogeneous. This assumption may never prevail in the strictest sense.

CONCLUSION

When respondents are asked to rank order many objects and/or respondents' ranking reliability may be in doubt, the researcher should consider using BIB designs. As long as the observations (even if non-numeric) are amenable to rank ordering according to some criterion of interest, the investigator may use BIB approaches. Because "balancing" and replication reduce variability, BIB procedures tend to increase the study's precision. With a sample taken from a large, heterogeneous population, there is a high probability of much variation across subjects' rankings. This is unlikely with a small sample drawn from a homogeneous population. Hence, a prerequisite for applying BIB methods is the existence of a high degree of commonality among respondents. The illustrated study satisfied this requirement by using top executives from the same corporation. Also, BIB methods are apt to save respondent time, cut interview costs, and result in increased completion rates. Since nonparametric statistical techniques have recently been associated with BIB procedures, the marketing researcher will not have to be a quantitative specialist to utilize it. Finally, with the rank order procedure, the researcher does not have to invent/create a scale of measurement he/she can use. This represents the *major* advantage and, therefore, significance of BIB methods.

REFERENCES

Banks, Seymour (1974), "Experimental Design and Control," in *Handbook of Marketing Research*, Robert Ferber, editor, New York: McGraw-Hill Book Company.

Bradley, Ralph and Milton Terry (1952), "Rank Analysis of Incomplete Block Designs," *Biometrika*, 39, 324-345.

Cochran, William and Gertrude Cox (1957), *Experimental Designs*, 2nd ed., New York: John Wiley & Sons, Inc.

Conover, W. (1971), *Practical Nonparametric Statistics*, New York: John Wiley & Sons, Inc.

Durbin, J. (1951), "Incomplete Blocks in Ranking Experiments," *British Journal of Psychology*, Statistical Section, 4, 85-90.

Federer, Walter (1955), *Experimental Design*, New York: The Macmillan Company.

Gibbons, Jean (1971), *Nonparametric Statistical Inference*, New York: McGraw-Hill Book Company.

Green, Paul E. (1974), "On the Design of Choice Experiments Involving Multi-Factor Alternatives," *Journal of Consumer Research*, 1 (September), 61-68.

__________ and Frank J. Carmone (1970), *Multidimensional Scaling and Related Techniques in Marketing Analysis*, Boston: Allyn and Bacon, Inc.

__________ and Donald Tull (1975), *Research for Marketing Decisions*, 3rd ed., Englewood Cliffs, NJ: Prentice-Hall, Inc.

Guttman, Louis (1946), "An Approach for Quantifying Paired Comparisons and Rank," *Annals of Mathematical Statistics*, 17, 144-163.

Kendall, Maurice (1955), *Rank Correlation Methods*, 2nd ed., New York: Hafner Publishing Company.

Raghavarao, Damaraju (1971), *Construction and Combinatorial Problems in Design of Experiments*, New York: John Wiley & Sons, Inc.

Torgerson, Warren S. (1958), *Theory and Methods of Scaling*, New York: John Wiley & Sons, Inc.

Wiley, James B. (1978), "BIBD: A Data Management Program for 'BIBD' Choice Data," *Journal of Marketing Research*, 15 (August), 472-474.

Yates, F. (1936), "Incomplete Randomized Blocks," *Annals of Eugenics*, 7, 121-140.

ABOUT THE AUTHORS

David R. Rink (Ph.D., University of Arkansas) is Professor of Marketing at the University of Indiana at Kokomo.

James E. Dunn (Ph.D., Virginia Polytechnic Institute and State University) is Professor of Mathematical Sciences and University Professor at the University of Arkansas.