"Balanced incomplete block designs: selected business-related applications and usage caveats"

NUMBER OF REFERENCES

0

NUMBER OF FIGURES

0

NUMBER OF TABLES

0

David R. Rink (USA)

Balanced incomplete block designs: selected business-related applications and usage caveats

Abstract

Whenever respondents must rank-order a large number of items and/or the reliability of their rankings may be questionable, balanced incomplete block designs (BIBDs) represent a more effective means for doing so than either complete rankings or paired comparisons for business and marketing researchers. By providing a type of balancing and replication across items and respondents, BIBDs significantly reduce the number of subjective evaluations each individual must make. But, at the same time, BIBDs allow a limited number of respondents as a group to rank many items. This balancing and replication in BIBDs also reduces standard deviation, which increases the precision of a study. BIBDs, therefore, can improve response rates as well as increase the accuracy and reliability of the data collected. After discussing the general nature of BIBDs and statistical techniques for analyzing preference data collected by BIBDs, three business-related applications are presented to illustrate the benefits of BIBDs. Next, caveats concerning the use of BIBDs are presented. In the last section, advantages of BIBDs are discussed.

Keywords: balanced incomplete block designs, Durbin test, coefficient of concordance, Guttman scaling, ranking many items, small sample sizes.
JEL Classification: M31.

Introduction

Business and marketing researchers oftentimes confront situations where they need to collect subjective or judgmental information from respondents. To further compound the situation, these individuals must rank-order objects in terms of preference or importance. In addition, there are many items to rank (e.g., 15). As a result, the ability of respondents to effectively and reliably rank-order these objects may be questionable. Finally, only a few qualified subjects, that are sufficiently knowledgeable, exist or are available (e.g., 10). Together, these factors elevate the researcher's task to one of Herculean proportions.

1. Purpose

The purpose of this paper is to provide business and marketing researchers with a mathematically accepted alternative for collecting subjective data from a small number of respondents, who must rank many items (e.g., product attributes) in order of importance or preference, which can result in higher response rates, reduce study costs, and increase the accuracy and reliability of the data. After discussing this method, with its accompanying statistical techniques, three real-world examples will be presented, which illustrate the application of this method. Next, caveats concerning the use of this approach will be provided. In the last section of the paper, advantages of this method will be discussed.

2. Development of the problem

As a means for determining the relative importance of a set of items, business and marketing researchers commonly ask respondents to rank these from most to least preferred. However, if the number of items to be ranked is relatively large (e.g., 15), and the number of respondents is limited (e.g., 10), it may be impractical for all items to be applied to each individual for comparison at the same time (e.g., Conover, 1999; Gisbrecht and Gumbertz, 2004). Otherwise, this may be more than respondents can effectively handle. Feeling overwhelmed, they may not assign a rank to each item.

The alternative typically used in lieu of complete rankings is paired comparisons. Here individuals indicate their preference for one item in successive pairs of items. While paired comparisons substantially reduce the number of items respondents must evaluate, it dramatically increases the number of pairs of items they must rate. Continuing the previous example of 15 items, paired comparisons would necessitate 105 individual sets of paired items to be evaluated. Such a laborious undertaking would likely tire and/or frustrate respondents[1]. In either case, the accuracy and reliability of respondents' answers may be questionable (e.g., Gibbons, 1971; Green et al., 1988; Green et al., 1989; Stinson, 2003). The ability of respondents, therefore, "to rank objects effectively and reliably may be a function of the number of comparative judgments to be made. For example, after 10 different brands of bourbon have been tasted, the discriminatory powers of the observers may legitimately be questioned" (Gibbons, 1971, p. 257).

---

[1] In addition, "deriving all pairs of a large number of items can be tedious and time-consuming" for the researcher (Rink, 1987, p. 54).

3. General nature of balanced incomplete block designs[2]. Several business and marketing researchers have found a more effective means for collecting judgmental data than complete rankings and paired comparisons. Labelled incomplete block design, this preference data collection method is especially suited to cases where the number of items is relatively large and the number of respondents is limited. In this approach, items are first divided into blocks according to a specific design. Then, the items in each block are presented to respondents for evaluation (e.g., Cox, 1958; Wu and Hamada, 2000; Aloke, 2010).

If the design is balanced such that each block contains a specified number of experimental units, each item appears in a certain number of blocks, and every item appears with every other item an equal number of times, then the design is called a balanced incomplete block design (BIBD) (e.g., Federer, 1955; Conover, 1971; Cochran and Cox, 1992; Conover, 1999; Colbourn and Dinitz, 2007). Every BIBD must satisfy these two defining relations:

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

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

where t = number of treatments (items) to be examined, b = total number of blocks (respondents), k = number of experimental units per block ($k < t$), r = number of times each treatment appears ($r < b$), and λ = number of blocks in which the ith treatment and the jth treatment appear together (λ is the same for all pairs of treatments).

BIBDs 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). Through balancing and replication of items and respondents, BIBDs reduce standard deviation. This, in turn, increases the precision of the study, thereby increasing the accuracy and reliability of respondents' answers (e.g., Green et al., 1988; Green et al., 1989; Colbourn and Dinitz, 2007).

4. Statistical analysis techniques for BIBD data

Several approaches have been suggested for analyzing data in BIBDs. The traditional technique for doing so has been analysis of variance. However, the specific form of analysis of variance for BIBDs "differs according to the nature of the design, the number of replications, and the restrictions" (Banks, 1974, p. 493)[3]. An analytical procedure that is computationally simpler than traditional analysis of variance techniques incorporates the Durbin test (e.g., Durbin, 1951; Conover, 1971; Hollander and Wolfe, 1999), coefficient of concordance (Kendall, 1955; Gibbons, 1971), and Guttman scale (Guttman, 1946).

4.1. Durbin test. Because observations in BIBDs consist merely of ranks that do not meet the normality assumption required for applying parametric techniques, the Durbin test is appropriate. Two key assumptions underlie the use of the Durbin test in BIBDs: the blocks (respondents) are mutually independent of each other; and, within each block, the treatments (items) may be arranged in increasing order according to some criterion of interest. The corresponding null hypothesis is the ranking of random variables within each block is equally likely. In other words, the treatments have identical effects. The alternative hypothesis states that at least one treatment tends to yield a larger observed value than at least one other treatment. The Durbin test statistic is defined in convenient computing form as (e.g., Conover, 1971; Hollander and Wolfe, 1999):

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

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 $(1-\alpha)^{th}$ quantile of a Chi-square random variable with $t-1$ degrees of freedom (e.g., Conover, 1971; Hollander and Wolfe, 1999).

4.2. Coefficient of concordance. The coefficient of concordance provides a measure of the degree of agreement among respondents regarding their rankings of items. The coefficient of concordance is defined in convenient computing form as (e.g., Kendall, 1955; Gibbons, 1971):

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

---

[2] Three of the earliest contributors in the construction of BIBDs were Yates (1936a; 1936b), Youden (1937), and Bose (1939). The extension of BIBDs from their origin in agricultural research to the behavioral sciences seems to have arisen through the methodology for paired comparisons (e.g., Bradley and Terry, 1952; David, 1988). For an excellent description of the early history of combinatorial designs and BIBDs, refer to Anderson et al. (2007).

[3] The analysis of variance methods for BIBDs may be found in several sources (e.g., Banks, 1974; Cochran and Cox, 1992; Stinson, 2003; Gisbrecht and Gumbertz, 2004) and will not be discussed in this paper.

The value of W is statistically significant at the same level as the Durbin test statistic, because the coefficient of concordance is simply a linear transformation of the Durbin test statistic.

4.3. Guttman scale[4]. If the null hypothesis of no difference among treatments is rejected, a Guttman scale can be developed. A Guttman scale is appropriate whenever an ordered set of statements exists, and agreement with one statement implies agreement with all statements that are less positive. This notion of agreement is best handled quantitatively by using scalogram analysis developed by Guttman (1946). Because the researcher is also interested in determining the intensity with which respondents rank-ordered objects, a Guttman scale can be derived directly from the rankings. The Guttman scale, therefore, is applicable to preference testing where the primary interest is in the objects under comparison (e.g., David, 1988; Cochran and Cox, 1992). After plotting the results of these calculations on a linear scale, the researcher can visually determine those items respondents considered most important, as well as ascertain the intensity with which they ranked these items.

5. Selected business-related applications

In the remainder of this paper, three business-related applications are presented to illustrate the benefits of BIBDs[5]. These applications range from dimensions for selecting real estate brokers to procurement strategies for each stage of a product's sales cycle to location sites for a distribution center.

5.1. Attributes used in the selection of real estate brokers. Green (1975) asked corporate real estate managers from several different large-sized companies in a major American Southern metropolitan area to identify and rank in order of importance attributes they used when selecting independent brokers. Because of qualitative considerations (e.g., real estate managers would have trouble ranking more than seven selection characteristics, and sample constraints associated with the pilot study), Green felt the most appropriate BIBD was the following: 11 selection dimensions (treatments) would be evaluated; 11 corporate real estate managers (blocks) would be sampled; each selection factor (treatment) would be repeated six times; each manager would rank-order six selection characteristics; and each selection factor would be compared with every other attribute by three managers. In ranking their specific set of six broker selection characteristics, corporate real estate managers were asked to assign the rank of "1" to the characteristic they felt was most important in selecting a broker, the rank of "2" to the second most important attribute, and so on until the rank of "6", which represented the least important dimension.

The results of the pilot survey among corporate real estate managers concerning broker selection criteria are summarized in Table 1. The first row of this table means Corporate Manager # 1 ranked his/her set of six real estate broker selection attributes from most to least important as follows: Dimension #s 2, 1, 5, 3, 6, and 4. And so on, until Corporate Manager # 11, whose ranking of his/her set of six broker selection characteristics was: Dimension #s 7, 1, 10, 4, 11, and 2.

After adding the ranks for each broker selection factor (or column) in Table 1, the Durbin test statistic was computed:

$$T = \frac{12(11-1)}{6(11)(6-1)(6+1)} \left[(19)^2 + (22)^2 + \ldots + (30)^2 \right]$$

$$- 3 \left[\frac{6(11-1)(6+1)}{(6-1)} \right] = 25.58.$$

Next, the coefficient of concordance was calculated:

$$W = \frac{12[(19)^2 + (22)^2 + \ldots + (30)^2] - 3(6)^2(11)(6+1)^2}{(3)^2(11)[(11)^2 - 1]} = 0.497.$$

At the .05 level, both were found to be significant. This meant there was a preferred order of broker selection characteristics among corporate real estate managers. In addition, these managers' rankings exhibited some degree of consistency. Since the rank of "1" signified the selection dimension each manager felt was most important while the rank of "6" indicated the least important, the attribute with the lowest sum of ranks would represent the charac-

---

[4] An alternative method to computing a Guttman scale is to rank item columnar sums in order of increasing magnitude. However, relative to the Guttman scale approach, this procedure is ad hoc in nature. Further, while the ranking of treatment columnar sums has strong theoretical support in the literature on complete rankings (e.g., Kendall, 1955; Siegel, 1956), it is not clear how well this technique applies to blocks of incomplete rankings. Therefore, the Guttman scale reflects more accurately the intensity of respondents' incomplete rankings of items. This will simplify the researcher's task of differentiating which items respondents judged of greatest importance.

[5] BIBDs can be applied to any situation where systematic comparisons are possible. To date, a plethora of applications of BIBDs have been made in the fields of agriculture, biology, engineering, medicine, physical and chemical sciences, communications systems, cryptology, business, education, healthcare, mathematics, pharmaceuticals, to name a few (e.g., Yang, 1985; Stinson, 2003; Van der Linden et al., 2004; Bose and Mukerjee, 2006; Dey, 2010; DeMuth, 2014). Some of more recent and specific applications of BIBDs include software testing, biological assay, medical clinical trials, sensory analysis, networking, quality control, image analysis, coding theory, bioenergy, algorithms and analysis, mathematical biology, signal processing, genetics, and industrial experimentation (e.g., Yang, 1985; Wakeling and Buck, 2001; Stinson, 2003; Van der Linden et al., 2004; Bose and Mukerjee, 2006; Chakrabarti, 2006; Camarda and Fiume, 2007; Camtepe and Yener, 2007; Louviere and Lancsar, 2009; Dey, 2010; Wang and Su, 2010; Kumar and Kaliyaperumal, 2012; Noshad and Brandt-Pearce, 2012; McClosky and Tanksley, 2013; Mohammad, 2013; William et al., 2013; Balakrishnan, 2014; Dogan et al., 2014; DeMuth, 2014; Casler et al., 2015; Civille and Carr, 2015; Muhlbacher et al., 2016).

teristic corporate real estate managers felt was most important in selecting a broker. As shown in Table 1, this dimension was the "broker's real estate and business experience" (Dimension # 7).

5.2. Procurement strategies by product life cycle stage. Berenson (1967) developed a list of 34 procurement strategies he recommended purchasing executives implement across six product life cycle (PLC) stages – introduction, growth, maturity, saturation, decline, and abandonment (Table 2). To test the validity of Berenson's model, Rink (1976) asked a group of purchasing executives from various major manufacturing firms in the Southwestern part of the U.S. to rank-order these 34 procurement strategies in terms of importance in the formulation of purchasing strategy for each stage in the PLC of one of their company's major products. After consulting a BIBD layout table in Raghavarao (1971), he decided on the basis of qualitative criteria (e.g., executives might find it unwieldy ranking nine items) to reduce the number of strategies to be considered for each PLC stage from 34 to 21. This was achieved by including only those procurement strategies from the first three PLC phases in the BIBD for the Introduction stage for purchasing executives to consider. It was felt managers would not consider strategies from the last three PLC phases in the Introduction stage. The same logic was applied to the Growth phase. However, in the BIBD for the Maturity stage, those strategies corresponding to the Introduction phase were replaced with those from the Saturation stage. The BIBDs for each of the last three PLC phases included procurement strategies from the last three PLC stages. Also, given time, cost, and sample size constraints, the number of blocks (executives) was limited to 30.

Having defined the number of strategies and sample size, Raghavarao's BIBD layout table was consulted again. Of the four qualifying designs, the one with the highest efficiency value (i.e., *tλ/rk*) was selected as the most appropriate design. In this particular BIBD, each purchasing executive would rank seven procurement strategies, and each strategy would be repeated 10 times. Each strategy would be compared with every other strategy by three executives. With the BIBD parameters defined, the specific design for each PLC stage was formulated (Table 3). The first row of Table 3 means Purchasing Executive # 1 would rank Procurement Strategy #s 1, 6, 9, 12, 17, 19, and 20 in order of importance from most to least important. And so on, until Purchasing Executive # 30 ranked Procurement Strategy #s 8, 9, 10, 11, 12, 13, and 14 from most to least important. Within this BIBD, three elements were randomized: order of strategies presented to each purchasing executive;

assignment of executives to blocks; and assignment of identification numbers to strategies. The survey was conducted through a combination of personal interview and paper-pencil methodology.

In the interests of brevity, only the results from the Introduction stage of the PLC will be presented with regards to purchasing executives' rankings of procurement strategies (Table 4). The first row of Table 4 means Purchasing Executive # 1 ranked his/her set of seven strategies in the Introduction Stage of the PLC from most to least important as follows: Procurement Strategy #s 5, 14, 16, 20, 1, 15, and 13. And so on, until Purchasing Executive # 30, whose ranking of his/her seven strategies was: Procurement Strategy #s 12, 10, 13, 14, 8, 11, and 9.

Following the summation of the ranks for each procurement strategy in the Introduction phase of the PLC, the Durbin test statistic was computed:

$$T = \frac{12(21-1)}{10(21)(7-1)(7+1)} \left[(33)^2 + (23)^2 + \ldots + \right.$$

$$\left. (50)^2 \right] - 3 \left[\frac{10(21-1)(7+1)}{(7-1)} \right] = 71.62.$$

Next, the coefficient of concordance was calculated:

$$W = \frac{12[(33)^2 + (23)^2 + \ldots + (50)^2] - 3(10)^2(21)(7+1)^2}{(3)^2(21)[(21)^2 - 1]}$$

$$= 0.434.$$

Both were found to be significant at the .05 level for the Introduction phase of the PLC (Table 5). At least one procurement strategy within the Introduction stage yielded a larger value than at least one other strategy. That is, executives viewed some procurement strategies as being more important than other strategies in the Introduction phase of the PLC. Also, purchasing executives were utilizing the same criterion in evaluating procurement strategies in the Introduction stage.

For each of the remaining PLC phases – growth, maturity, saturation, decline, and abandonment, the Durbin test statistic and coefficient of concordance were both found to be significant at the .05 level (Table 5). This meant at least one purchasing strategy in each of these stages yielded a larger value than at least one other strategy. That is, for each of these remaining PLC phases, managers perceived some procurement strategies as being more important than other strategies. In addition, these purchasing executives were using the same criterion in rating procurement strategies in each of these PLC stages.

As a means for determining the order and intensity with which executives ranked the 21 strategies in each PLC phase, a Guttman score was derived for

each strategy (Table 6). If the most important procurement strategy received a rank of "1" and the least important strategy was accorded a rank of "7", then, those procurement strategies possessing the highest negative Guttman scores would represent the most important strategies for these purchasing executives. Next, these Guttman scores were plotted on a linear scale for each PLC stage.

Again, in the interests of brevity, only the results from the Introduction stage of the PLC will be shown. In Table 4, PS # 3 was rated more important in terms of sum of ranks by executives than PS # 2. But, the opposite was true with regard to Guttman scores. Several similar instances occurred (e.g., PS #s 5 and 14). In addition, some pairs of procurement strategies in Table 4 had the same sum of ranks (e.g., PS #s 1 and 6). The strategies in each pair, however, had different Guttman scores. This phenomenon transpired because sum of ranks is merely a summation of the ranks for each procurement strategy, while Guttman scores reflect the configuration of ranks within each strategy. Thus, Guttman scaling affords a more accurate picture of the actual order and distribution of purchasing strategies than sum of ranks. Of the 21 procurement strategies evaluated by purchasing executives for the Introduction phase of the PLC, these six strategies were rated most important (in order): PS #s 2, 3, 14, 5, 10, and 4 (Table 6 and Figure 1). Four of these procurement strategies (i.e., PS #s 2, 3, 5, and 4) coincided with the six strategies recommended by Berenson (1967) for the Introduction stage of the PLC (Table 2).

5.3. Location for a new distribution center. A large multi-national conglomerate wanted to determine where to locate a new distribution center. After gathering preliminary information on 14 potential sites, the CEO eliminated three from further consideration. For long-term strategic planning purposes, the CEO wanted his top managers to evaluate and rank these 11 possible locations in order of preference. Realizing these executives may have difficulty doing so, he sought the assistance of a consultant (Rink, 2006) to make this task more manageable. This individual recommended the CEO use a BIBD.

Consulting a BIBD layout table (e.g., Raghavarao and Padgett, 2005) for 11 potential sites (treatments), four possible BIBDs emerged. Selection of the optimal BIBD depended upon both quantitative criteria (i.e., highest efficiency value, or $t\lambda/rk$) and qualitative considerations (e.g., number of objects respondents could reliably rank). The CEO felt top management could rank six locations more accurately than seven. As there were 11 executives, the CEO was constrained to a BIBD with 11 blocks. As a result, the consultant was able to select the optimal BIBD. Each site (or treatment) would be repeated six times, and each location would be compared with every other location by three top managers.

With the establishment of these parameters, the consultant was able to develop the BIBD layout (Table 7). Specifically, the first row of this table means Manager # 1 would rank Location Site #s 1, 2, 3, 4, 5, and 6 in order of preference from the most to the least preferred. And so on, until Manager # 11 ranked Location Site #s 1, 2, 4, 7, 10, and 11. Within this BIBD, the consultant randomized three elements: order of distribution center location numbers presented to each executive to be ranked; assignment of top managers to blocks; and assignment of identification numbers to location sites.

After conducting the survey, the consultant obtained the results summarized in Table 8. The first row of this table means Manager # 1 ranked his/her set of six distribution center locations from most preferred to least preferred as follows: Location Site #s 1, 3, 5, 2, 4, and 6. And so on, until Manager # 11, whose ranking of his/her set of six sites was: Location Site #s 7, 1, 10, 4, 2, and 11. After summing the ranks for each distribution center, the Durbin statistic was computed:

$$T = \frac{12(11-1)}{6(11)(6-1)(6+1)} \left[(15)^2 + (26)^2 + \ldots + \right.$$

$$\left. (31)^2 \right] - 3 \left[\frac{6(11-1)(6+1)}{(6-1)} \right] = 35.05.$$

Next, the coefficient of concordance was calculated:

$$W = \frac{12[(15)^2 + (26)^2 + \ldots + (31)^2] - 3(6)^2(11)(6+1)^2}{(3)^2(11)[(11)^2 - 1]}$$

$$= 0.681.$$

At the .05 level, both results were found to be significant. This meant at least one location site tended to yield a larger observed value than at least one other site. Hence, there appeared to exist at least a partial ordering of location sites among top management. Also, the rankings of distribution center locations by executives exhibited some degree of consistency. In other words, these managers were using the same criterion in evaluating location sites.

As a means for determining the order and magnitude with which top management ranked the 11 distribution center locations, the consultant derived a Guttman score for each location site, which is shown in the last row of Table 8. Since the most preferred site received a rank of "1" while the least preferred location site was awarded a rank of "6" by each executive, then, those sites possessing the most negative Guttman scores would represent the most preferred

location sites across all managers. Next, these Guttman scores were plotted on a linear scale (Figure 2). On the basis of sum of ranks, Site # 11 was rated more preferred than Site # 8; but, in terms of Guttman scores, the opposite was true (Table 8). In addition, two distribution center location sites had the same sum of ranks (i.e., Site #s 1 and 9); however, they had different Guttman scores (Table 8). This phenomenon transpired because Guttman scores reflect the configuration of ranks among each location site while the sum of ranks does not. Hence, Guttman scaling provides a more detailed and accurate picture of the actual order and spread among distribution center location sites than sum of ranks. In the present case, executives overwhelmingly rated Site # 10 as the most preferred location site for the new distribution center. Although Site # 7 was a distant second, the consultant recommended the company retain it in case negotiations for Site # 10 deteriorated.

6. Caveats concerning BIBDs

Before using BIBDs, several caveats warrant attention. Items being considered for ranking must be amenable to rank-ordering according to some criterion of interest. Usually, this is not a major problem. A prerequisite for applying BIBDs is the existence of a high degree of commonality among respondents. Otherwise, there is likely to be much variation across individuals' rankings. This potential problem was averted in each of the previously described applications. In the first example, respondents were corporate real estate managers from different large-sized companies in a major American Southern metropolitan area. In the second case, respondents were purchasing executives from various major manufacturers in the Southwestern part of the U.S. But, in the last application, all of the respondents were top-level executives from the same large multi-national conglomerate.

When consulting a BIBD layout table (e.g., Raghavarao and Padgett, 2005), several possible BIBDs may emerge. In order to objectively determine the optimal BIBD, first, compute each design's efficiency factor (i.e., $t\lambda/rk$)[6], and second, select the BIBD with the largest value. However, in some instances, the researcher has to consider qualitative factors (e.g., respondents may have difficulty ranking nine product attributes) in selecting the "best" BIBD. This is exactly what occurred in the second application before the number of purchasing strategies to be ranked was reduced from 34 to 21. Other things being equal, λ should be greater than or equal

to two, because each treatment is, then, compared at least twice with every other treatment rather than once if $\lambda = 1$. Also, the efficiency factor will be higher in the former case than in the latter. Both of these recommendations were followed by Rink (1987).

Once the number of respondents has been determined in the optimal BIBD, this exact number – no more, no less – must be obtained. This may necessitate some form of personal interview procedure, which can be time-consuming and expensive, as both Green (1975) and Rink (1976) discovered. Following the selection of the "optimal" BIBD layout and prior to data collection, these three things should be randomized: (1) order of items presented to each respondent; (2) assignment of respondents to each set of items to be ranked; and (3) assignment of identification numbers to items.

After the requisite rank-order data have been collected, relying solely upon sum of ranks to identify the most preferred item may lead to an erroneous conclusion, especially if two items have the same columnar sum of ranks. Because Guttman scaling incorporates the configuration of ranks within each item instead of simply summation of ranks, Guttman scores more accurately reflect the intensity with which respondents ranked items, thereby making it easier to correctly ascertain the most preferred item (Guttman, 1946). For example, in the second application, four different sets of two purchasing strategies (e.g., PS #s 11 and 19) had the same columnar sum of ranks; however, they had different Guttman scores. Finally, while a significant coefficient of concordance means individuals applied the same criterion in ranking items, it "does not mean that the orderings observed are correct. In fact, they may all be incorrect with respect to some external criterion" (Siegel, 1956, p. 238). Replicating a study using "similar" respondents is one way to generate the required external criterion for subsequent investigations.

Summary and conclusions

Whenever respondents are asked to rank a large number of items and/or the reliability of their rankings may be questionable, balanced incomplete block designs (BIBDs) should be considered. They are relatively easy to construct and analyze. BIBDs substantially reduce the number of items each individual must subjectively evaluate. Through balancing and replication of items and respondents, a small group of individuals is able to rank many items. Because balancing and replication reduce standard deviation, BIBDs also increase the precision of a study, even with a small sample. If the population is homogeneous, then, a small sample will likely result in more valid inferences than one selected from a large, heterogeneous population.

---

[6] The efficiency factor (i.e., $t\lambda/rk$) is defined as "the fraction of total information contained in intrablock comparisons where interblock and intrablock contrasts are of equal accuracy" (Federer, 1955, p. 416).

Since data collected by BIBDs are ordinal scale in nature, nonparametric statistical techniques may be used. Nonparametric statistics are also appropriate when sample sizes are small, which oftentimes is the case with business-related situations. Because nonparametric statistics are "distribution-free", they are easier to learn and use than their parametric counterparts. By incorporating Guttman scaling in BIBDs (and plotting Guttman scores on a linear scale), the intensity of individuals' rankings can be accurately and readily determined, which will simplify the identification of items rated most important. Moreover, with the rank-order procedure, a scale of measurement does not need to be invented. Finally, from an administration standpoint, since each respondent is ranking a subset of the total number of items, BIBDs will save respondents' time, which can result in higher response rates, reduce study costs, and increase the accuracy and reliability of the data.

References

1. Aloke, D. (2010). *Incomplete Block Designs*. Singapore: World Scientific.
2. Anderson, I., C. Colbourn, J. Dinitz and T. Griggs. (2007). *Design theory: Antiquity to 1950, Handbook of Combinatorial Designs*, 2nd ed., C. Colbourn and J. Dinitz, eds., Boca Raton, FL: CRC Press.
3. Berenson, C. (1967). The purchasing executive's adaptation to the product life cycle, *Journal of Purchasing and Materials Management*, 3, May, pp. 52-68.
4. Banks, S. (1974). *Experimental Design Control, Handbook of Marketing Research*. R. Ferber, ed., New York: McGraw-Hill.
5. Bose, M. and R. Mukerjee. (2006). Optimal visual cryptographic schemes, *Designs, Codes and Cryptography*, 40, 3, pp. 255-267.
6. Bose, R. (1939). On the construction of balanced incomplete block designs, *Annuals of Eugenics*, 9, June, pp. 353-399.
7. Balakrishnan, N. (2014). *Methods and Applications of Statistics in Clinical Trials*. New York: Wiley.
8. Bradley, R. and M. Terry. (1952). Rank analysis of incomplete block designs, *Biometrika*, 39, January, pp. 324-345.
9. Camarda, P. and O. Fiume. (2007). Collision free MAC protocols for wireless ad hoc networks based on BIBD architecture, *Journal of Commuications*, 2 (7), pp. 1-8.
10. Camtepe, S. and B. Yener. (2007). Combinatorial design of key distribution mechanisms for wireless sensor networks, *IEEE/ACM Transactions on Networking*, 15 (2), pp. 346-358.
11. Casler, M., W. Vermerris and R. Dixon. (2015). Replication concepts for bioenergy, *Research Experiments*, 8, March, pp. 1-6.
12. Chakrabarti, D. (2006). A new visual cryptography scheme for STS based access structures, *Journal of Discrete Mathematical Sciences and Cryptography*, 9 (1), pp. 9-23.
13. Civille, G. and B. Carr. (2015). *Sensory Evaluation Techniques*. 5th ed., Boca Raton, FL: CRC Press.
14. Cochran, W. and G. Cox. (1957). *Experimental Designs*. 2nd ed., New York: Wiley.
15. Cochran, W. and G. Cox. (1992). *Experimental Designs*. Revised ed., New York: Wiley.
16. Colburn, C. and J. Dinitz. (2007). *Handbook of Combinatorial Designs*. 2nd ed., Boca Raton, FL: Chapman and Hall.
17. Conover, W. (1971). *Practical Nonparametric Statistics*. New York: Wiley.
18. Conover, W. (1999). *Practical Nonparametric Statistics*. 3rd ed., New York: Wiley.
19. Cox, D. (1958). *Planning of Experiments*. New York: Wiley.
20. David, H. (1988). *The Method of Paired Comparisons*. 2nd ed., London: Griffin.
21. DeMuth, J. (2014). *Basic Statistics and Pharmaceutical Statistical Applications*. 3rd ed., Boca Raton, FL: CRC Press.
22. Dey, A. (2010). *Incomplete Block Designs*. New Delhi, India: Hindustan Book Agency.
23. Dogan, A., S. Ors and G. Saldamli. (2014). Analyzing and comparing the AES architectures for their power consumption, *Journal of Intelligent Manufacturing*, 25, April, pp. 263-271.
24. Durbin, J. (1951). Incomplete blocks in ranking experiments, *British Journal of Psychology*, 4, November, pp. 85-90.
25. Federer, W. (1955). *Experimental Design*, New York: Macmillan.
26. Gibbons, J. (1971). *Nonparametric Statistical Inference*, New York: McGraw-Hill.
27. Gisbrecht, F. and M. Gumbertz. (2004). *Planning, Construction, and Statistical Analysis of Comparative Experiments*, New York: Wiley.
28. Green, H. (1975). *Purchasing professional services: A linear compensatory approach to weighting the dimensions for evaluating real estate brokers and appraisers*, an unpublished Ph.D. dissertation, University of Arkansas.
29. Green, P. (1974). On the design of choice experiments involving multi-factor alternatives, *Journal of Consumer Research*, 1, September, pp. 61-68.
30. Green, P., F. Carmone, Jr. and S. Smith. (1989). *Multidimensional Scaling and Related Techniques in Marketing Analysis*, Boston: Allyn and Bacon.
31. Green, P., D. Tull and G. Albaum. (1988). *Research for Marketing Decisions*. 5th ed., Englewood Cliffs, NJ: Prentice-Hall.
32. Guttman, L. (1946). An approach for quantifying paired comparisons and rank, *Annals of Mathematical Statistics*, 17, August, pp. 144-163.

33. Hollander, M. and D. Wolfe. (1999). *Nonparametric Statistical Methods*. 2nd ed., New York: Wiley.
34. Kendall, M. (1955). *Rank Correlation Methods*. 2nd ed., New York: Hafner.
35. Kumar, R. and G. Kaliyaperumal. (2012). Optimal fingerprint scheme for video on demand using block designs, *Multi-media Tools and Applications*, 61, November, pp. 389-418.
36. Louviere, J. and E. Lancsar. (2009). Choice experiments in health, *Health Economics, Policy and Law*, 4 (4), pp. 527-546.
37. McClosky, B. and S. Tanksley. (2013). Optimizing experimental design in genetics, *Journal of Optimization Theory and Applications*, 157, May, pp. 520-532.
38. Mohammad, N. (2013). *Applications of balanced incomplete block designs to communications systems*, an unpublished Ph.D. dissertation, University of Virginia.
39. Muhlbacher, A., A. Kaczynski, P. Zweifel and F. Johnson. (2016). Experimental measurement of preferences in health and healthcare using best-worst scaling, *Health Economic Review*, 6, January, pp. 1-14.
40. Neubauer, D. (2010). Design of comparative experiments, *Technometrics*, 52, May, pp. 261-263.
41. Noshad, M. and M. Brandt-Pearce. (2012). Expurgated PPM using symmetric balanced incomplete block designs, *IEEE Communications Letters*, 16, July, pp. 968-971.
42. Raghavarao, D. (1971). *Construction and Combinatorial Problems in Design of Experiments*. New York: Wiley.
43. Raghavarao, D. and L. Padgett. (2005). *Block Designs: Analysis, Combinations, and Applications*. Singapore: World Scientific.
44. Rink, D. (1976). *An empirical investigation of the impact of the product life cycle theory upon the formulation of purchasing strategies*, an unpublished Ph.D. dissertation, University of Arkansas.
45. Rink, D. (1987). An improved preference data collection method: Balanced incomplete block designs, *Journal of the Academy of Marketing Science*, 15, Spring, pp. 53-61.
46. Rink, D. (2006). *Another real-world application of BIBDs: Distribution Center Location*, a working paper.
47. Sharma, P. and S. Kumar. (2014). Some applications of Hadamard Rhotrices to design balanced incomplete block, *International Journal of Mathematical Science and Engineering Applications*, 8, March, pp. 389-404.
48. Siegel, S. (1956). *Nonparametric Statistics*. New York: McGraw-Hill.
49. Stinson, D. (2003). *Combinatorial Designs: Constructions and Analysis*. New York: Springer.
50. Van der Linden, W., B. Veldkamp, and J. Carlson. (2004). Optimizing balanced incomplete block designs for educational assessments, *Applied Psychological Measurement*, 28 (5), pp. 317-331.
51. Wakeling, I. and D. Buck. (2001). Balanced incomplete block designs useful for consumer experimentation, *Food Quality and Preference*, 12, pp. 265-268.
52. Wang, J. and R. Su. (2010). Further results on the existence of splitting BIBDs and application to authentication codes, *Acta Applicandae Mathematicae*, 109, March, pp. 791-803.
53. William, L., C. Nachtsheim, K. Wang, R. Reul and M. Albrecht. (2013). Conjoint analysis and discrete choice experiments for quality improvement, *Journal of Quality Technology*, 45, January, pp. 74-99.
54. Wu, H. and M. Hamada. (2000). *Experiments: Planning, Analysis and Parameter Design Optimization*. New York: Wiley.
55. Yang, J. (1985). An example of the application of a balanced incomplete block design, *Assessment and Evaluation in Higher Education*, 10 (3), pp. 250-253.
56. Yates, F. (1936a). Incomplete randomized blocks, *Annuals of Eugenics*, 7, March, pp. 121-140.
57. Yates, F. (1936b). A new method of arranging variety trials involving a large number of varieties, *Journal of Agricultural Science*, 26, June, pp. 424-455.
58. Youden, W. (1937). Use of incomplete block replications in estimating tobacco-mosaic virus, *Contributions from Boyce Thompson Institute*, 9, November, pp. 41-48.

Appendices

Table 1. Ranking of dimensions for selecting real estate brokers

Corporate managers	Dimensions for selecting real estate brokers[a]										
	1	2	3	4	5	6	7	8	9	10	11
1	2	1	4	6	3	5					
2	3	5	4				1	6	2		
3		5		2		4		6	1	3	
4			3	6	5		4		2	1	
5	5				3	2	1	6		4	
6	6			5	2			3	1		4
7	1		3			5			4	2	6
8			1	4		3	2	6			5
9		2			3	4	1		5		6

Table 1 (cont.). Ranking of dimensions for selecting real estate brokers

Corporate managers	Dimensions for selecting real estate brokers[a]										
	1	2	3	4	5	6	7	8	9	10	11
10		3	2		5			6		1	4
11	2	6		4			1			3	5
R_j's	19	22	17	27	21	23	10	33	15	14	30

Notes: [a] Dimensions for selecting real estate brokers are coded as follows (Green, 1975): 1 = Broker's professional affiliation and/or achievement; 2 = Broker's favorable position in the industry; 3 = Reference by a third party and/or previous contact with the broker; 4 = Broker's flexibility to customer's needs; 5 = Broker's geographical location and/or convenience factors; 6 = Broker's marketing innovativeness; 7 = Broker's real estate and business experience; 8 = Commission to/broker quality ratio; 9 = Broker's general reputation; 10 = Broker's knowledge as a source of information; and 11 = Auxiliary services offered by the broker.

Table 2. Berenson's product life cycle-procurement strategy model

Number	Procurement strategy
	Introduction stage
1	Maintain low inventory for materials purchased for the product
2	Seek knowledgeable suppliers which can provide technical service
3	Obtain suppliers with the potential to grow with the product but who are patient enough to supply small quantities
4	Seek suppliers who are flexible enough to vary their products and services as rapidly changing needs might require
5	Work closely with the pilot plant and marketing group – good communication is vital
6	Pay careful attention to quality control as customers have not yet formulated firm opinions as to the quality of the product
	Growth stage
7	Provide the goods and services needed in ever-increasing quantities
8	Develop new sources of supply to help assure the provision of ever-increasing quantities of needed goods and services
9	Seek suppliers who are geographically dispersed as more than one production location is likely for the product
10	Recognize the need for frequent and rapid deliveries and make provisions for this
11	Pay careful attention to quality control due to the rapid quantity build-up
12	Obtain information about many new suppliers entering the field
13	Build larger inventories once the product's future is more assured
14	Obtain data on and samples of many new grades and types of raw materials offered by the trade as the market grows
	Maturity stage
15	Obtain favorable long-term supply contracts now that the firm's plants are in stable operation
16	Emphasize low prices as well as high quality when buying
17	Help to standardize the grades, shipping customs, etc. of the trade
18	Weed out weak suppliers and remain alert to future weaknesses in other suppliers
19	Be cautious of the low-price merchandisers
20	Seek raw materials that will help increase yields and conversion
21	Emphasize cost reduction in the operation of the purchasing department
	Saturation stage
22	Balance purchased goods inventory levels very carefully so that the absolute minimum investment is made in this area
23	Make the best of the ruthless and rampant price cutting that will exist
24	Make the best of other characteristics of the buyer's market in which the purchasing executive will find himself (e.g., capitalize on the product differentiation and packaging innovations of suppliers)
25	Be prepared to deal with marketing-oriented rather than production-oriented suppliers
26	Be alert to gradual erosion of quality in the goods received
27	Keep looking for low-cost substitute goods and services
	Decline stage
28	Activities which have only a long-term effect should be eliminated (e.g., making long-term forecasts of suppliers' production capacity and labor costs)
29	Out-of-pocket expenses should be avoided
30	Re-examine and revise required inventory levels
	Abandonment stage
31	Be prepared to dispose of surplus materials that will no longer be needed
32	Be prepared to assume purchasing responsibilities for new or replacement items in the firm's line
33	Be ready to recommend alternatives that will avoid the necessity for dropping the product (e.g., the firm stops manufacturing certain materials or products, and instead acts only to resell such items produced by other organizations)
34	Give sufficient warning to the firm's suppliers so that they are prepared for the loss of the purchasing department's business

Table 3. General BIBD for procurement strategies in each product life cycle stage

Blocks (purchasing executives)	Treatments (procurement strategies)						
1	1	6	9	12	17	19	20
2	2	7	10	13	18	20	21
3	3	1	11	14	19	21	15
4	4	2	12	8	20	15	16
5	5	3	13	9	21	16	17
6	6	4	14	10	15	17	18
7	7	5	8	11	16	18	19
8	1	4	9	10	16	18	19
9	2	5	10	11	17	19	20
10	3	6	11	12	18	20	21
11	4	7	12	13	19	21	15
12	5	1	13	14	20	15	16
13	6	2	14	8	21	16	17
14	7	3	8	9	15	17	18
15	1	2	11	13	15	17	18
16	2	3	12	14	16	18	19
17	3	4	13	8	17	19	20
18	4	5	14	9	18	20	21
19	5	6	8	10	19	21	15
20	6	7	9	11	20	15	16
21	7	1	10	12	21	16	17
22	1	2	4	8	9	11	21
23	2	3	5	9	10	12	15
24	3	4	6	10	11	13	16
25	4	5	7	11	12	14	17
26	5	6	1	12	13	8	18
27	6	7	2	13	14	9	19
28	7	1	3	14	8	10	20
29	1	2	3	4	5	6	7
30	8	9	10	11	12	13	14

Table 4. Results from BIBD for procurement strategies for the introduction stage of the product life cycle

Executive number	Procurement strategy number																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	5				1								7	2	6	3				4	
2	3		1								6			2	5				7		4
3		2				4	3		6				7	1					5		
4	2						6			5		3					4	1			7
5			1				3	2	6						4		5	7			
6	2		1				5	7		3				6						4	
7					1	3		5		2					6				4		7
8				3			6					1	5		7				2		4
9						2	5		4		3				6	1				7	
10		1	2		5				6	3		4			7						
11		2	3					4				1			7	5				6	
12		4			3					5	6						1		7	2	
13	3	1		2				5	7		6										4
14	1	2									3		5		6		7	4			
15			2		1				6				5			3	7				4
16		5					6			2			3					7		1	4
17					7		5	1			6				4			3	2		
18			4	2		6				5	1		7			3					
19	7			2	1			4				3	5					6			
20	1					3			6			7					5		4	2	
21			2			5					6	1						7		4	3
22				3	1		4				2	6		5			7				
23			1	2				6					7				5		3	4	
24		1				3		5						2		4	7				6
25				3		5				2				1	6		7	4			
26		1	2										4	3		5		6	7		
27	3	4	5	6	2	1	7														
28	6			2					7	1						3		5	4		
29				5	3				4						1			6		2	7
30								5	7	2	6	1	3	4							
R_j's	33	23	21	31	26	33	50	44	59	30	45	31	54	27	60	35	52	55	45	36	50
Guttman scores	-.044	-.377	-.342	-.171	-.246	-.037	-.027	.129	.325	-.186	.009	-.110	.258	-.279	.348	-.051	.395	.189	.089	.030	.097

Table 5. Summary of statistical results from BIBD for procurement strategies for each PLC stage

Statistic	Introduction	Growth	Maturity	Saturation	Decline	Abandonment
Durbin test statistic	71.62	59.72	31.91	49.86	69.57	81.24
Coefficient of concordance	.434	.362	.193	.302	.422	.492
Probability of obtaining above values by chance	.00000018	.00000805	.04432094	.00023228	.00000030	0.0

Table 6. Ranking of procurement strategies by Guttman scores[a] by product life cycle stage

Order of strategies	Introduction	Growth	Maturity	Saturation	Decline	Abandonment
1	2 (-.377)[b]	2 (-.412)	20 (-.311)[b]	30 (-.453)	33 (-.389)	31 (-.376)[b]
2	3 (-.342)[b]	11 (-.250)[b]	16 (-.286)[b]	22 (-.300)[b]	30 (-.298)[b]	34 (-.310)[b]
3	14 (-.279)	16 (-.223)	7 (-.225)	33 (-.270)	32 (-.266)	22 (-.259)
4	5 (-.246)[b]	3 (-.220)	10 (-.213)	16 (-.239)	22 (-.193)	29 (-.239)
5	10 (-.186)	10 (-.182)[b]	8 (-.194)	27 (-.119)[b]	28 (-.177)	30 (-.175)
6	4 (-.171)[b]	4 (-.148)	15 (-.122)[b]	26 (-.110)[b]	16 (-.150)	33 (-.159)
7	12 (-.110)	12 (-.096)[b]	17 (-.105)[b]	29 (-.0894)	27 (-.132)	28 (-.132)
8	16 (-.051)	14 (-.0922)[b]	11 (-.099)	24 (-.0886)	29 (-.117)	21 (-.124)
9	1 (-.044)	7 (-.0919)	26 (-.042)	25 (-.061)	23 (-.049)	32 (-.060)
10	6 (-.037)	15 (-.053)	27 (-.016)	20 (-.014)	20 (-.034)	23 (.001)
11	7 (-.027)	19 (-.024)	19 (.003)	10 (.004)	34 (-.026)	27 (.013)
12	11 (.009)	6 (-.017)	24 (.012)	19 (.027)	18 (.021)	26 (.024)
13	20 (.030)	8 (.019)	25 (.041)	21 (.030)	21 (.051)	18 (.031)
14	19 (.089)	20 (.046)	21 (.068)	32 (.037)	31 (.056)	25 (.0976)
15	21 (.097)	13 (.088)	13 (.089)	31 (.063)	24 (.128)	24 (.0993)
16	8 (.129)	5 (.100)	9 (.0907)	28 (.071)	26 (.130)	20 (.144)
17	18 (.189)	21 (.122)	18 (.0911)	18 (.135)	10 (.155)	17 (.146)
18	13 (.258)	17 (.250)	14 (.103)	34 (.210)	17 (.157)	16 (.184)
19	9 (.325)	18 (.285)	12 (.117)	23 (.328)	25 (.288)	19 (.211)
20	15 (.348)	9 (.428)	22 (.278)	17 (.396)	19 (.294)	10 (.353)
21	17 (.395)	1 (.472)	23 (.372)	15 (.440)	15 (.552)	15 (.530)

Notes: [a] The entry in each cell represents a procurement strategy number corresponding to Table 2 while the decimal figure in parentheses symbolizes the Guttman score for that strategy. [b] Purchasing executives' evaluation of this procurement strategy coincided with Berenson's model (Table 2).

Table 7. General BIBD for ranking of distribution center locations

Blocks (top management)	Treatments (distribution center locations)					
1	1	2	3	4	5	6
2	1	2	3	7	8	9
3	2	4	6	8	9	10
4	3	4	5	7	9	10
5	1	5	6	7	8	10
6	1	4	5	8	9	11
7	1	3	6	9	10	11
8	3	4	6	7	8	11
9	2	5	6	7	9	11
10	2	3	5	8	10	11
11	1	2	4	7	10	11

Table 8. Results from BIBD for distribution center location preference study

Top mgmt. number	Distribution center location number										
	1	2	3	4	5	6	7	8	9	10	11
1	1	4	2	5	3	6					
2	6	4	3				1	5	2		
3		5		3		4		6	2	1	
4			5	6	4		2		3	1	
5	1				4	5	3	6		2	
6	3			2	4			5	1		6

Table 8 (cont.). Results from BIBD for distribution center location preference study

Top mgmt. number	Distribution center location number										
	1	2	3	4	5	6	7	8	9	10	11
7	2		3			5			4	1	6
8			1	5		6	2	4			3
9		5			1	4	2		3		6
10		3	2		5			6		1	4
11	2	5		4			1			3	6
R_j's	15	26	16	25	21	30	11	32	15	9	31
Guttman scores	-.123	.195	-.104	.176	.098	.281	-.237	.320	-.169	-.443	.338

.395 — Strategy # 17
.348 — Strategy # 15
.325 — Strategy # 9
.258 — Strategy # 13
.189 — Strategy # 18
.129 — Strategy # 8
.097 — Strategy # 21
.089 — Strategy # 19
.030 — Strategy # 20
.009 — Strategy # 11
-.027 — Strategy # 7
-.037 — Strategy # 6
-.044 — Strategy # 1
-.051 — Strategy # 16
-.110 — Strategy # 12
-.171 — Strategy # 4
-.186 — Strategy # 10
-.246 — Strategy # 5
-.279 — Strategy # 14
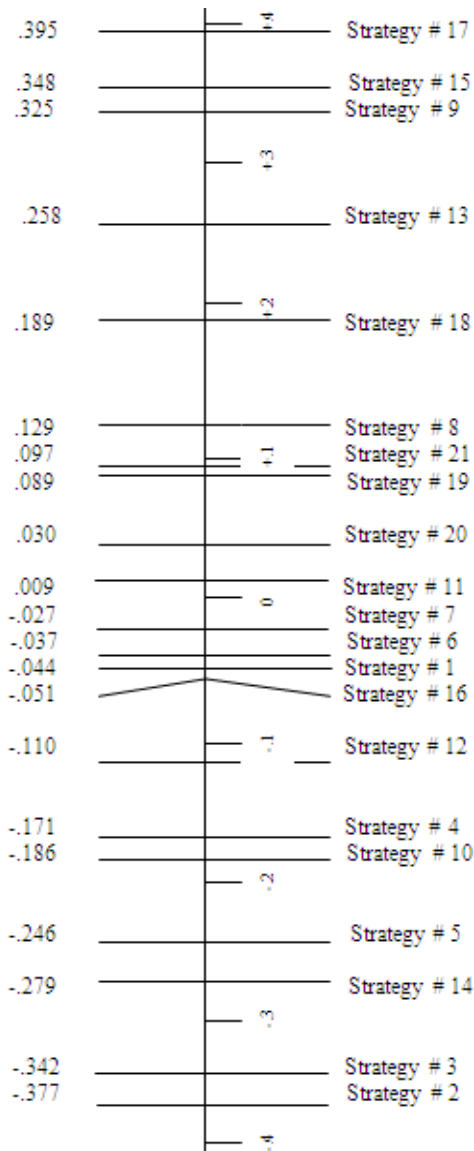-.342 — Strategy # 3
-.377 — Strategy # 2

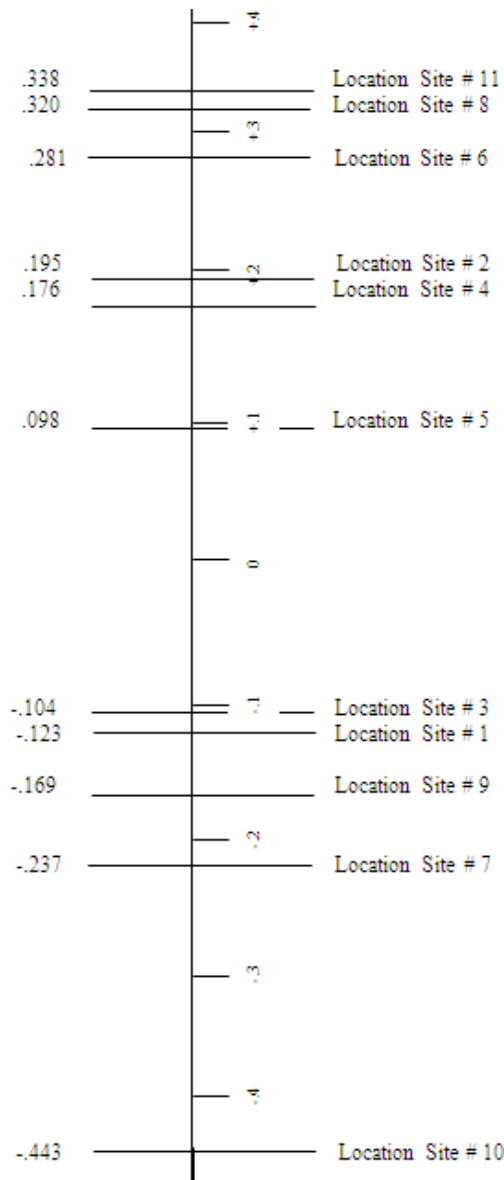**Fig. 1. Linear scale depicting Guttman scores for procurement strategies for the introduction stage**

Fig. 2. Linear scale depicting Guttman scores for distribution center location preference study