Automated assembly of optimally spaced and balanced paired comparisons: Controlling order effects

Mark L. Davison · Andrew Zieffler · Julio Cabrera ·
Stacy R. Karl · Harvey S. Cohen

Published online: 17 November 2011
© Psychonomic Society, Inc. 2011

Abstract To control order effects in questionnaires containing paired comparisons, Ross (1934) described an optimal ordering of the pairings. The pairs can also be balanced so that every stimulus appears equal numbers of times as the first and the second member of a pair. First, we describe and illustrate the optimally spaced, balanced ordering of pairings. Then we show how the optimally spaced, balanced order can be used to implement a matrix-sampling design or a fully incomplete design when the number of stimuli n is so large that respondents cannot reasonably be expected to judge all $n(n-1)/2$ pairs. The algorithm for balancing and optimally spacing the list of pairs is described.

Keywords Order effects · Paired comparisons · Forced choice tasks · Matrix sampling · Planned missingness · Questionnaire design · Preference judgments · Similarity judgments · Multidimensional scaling

The use of paired comparisons was popularized as a method of eliciting human judgments during the early 20th century by Thurstone (1927a, 1927b). McGuire and Davison (1991) enhanced the potential utility of paired comparisons by showing how paired comparisons could be used to test hypotheses about differences between experimental treatment groups or naturally occurring groups (e.g., males vs. females) using logistic or probit regression. Bockenholt

(2001, 2004) and Maydeu-Olivares and Bockenholt (2005) proposed a multilevel, random coefficient generalization of McGuire and Davison's single-level, fixed-effects logistic/probit regression models as a way of further enhancing the utility of paired-comparison data for testing experimental hypotheses. Before one can apply any of these analyses, however, one must obtain judgments of pairs so as to control order effects that can arise if the pairs are not optimally presented. In this article, we describe methods for controlling the order effects for both complete and incomplete designs, where an *incomplete design* is defined as a design with planned missingness in which each respondent judges only a fraction of all possible stimulus pairs.

There are at least four types of paired-comparison judgments: similarity judgments, preference judgments, judgments of agreement, and judgments of magnitude. Applications in which respondents make similarity judgments are frequently used in multidimensional scaling studies (e.g., Davison & Jones, 1976; Schubsachs & Davison, 1979). For example, the respondent may be shown pairs of stimuli (e.g., pairs of color chips varying in hue, saturation, and brightness) and asked, for each pair, to rate the chips' similarity. Alternatively, the researcher may present pairs of stimuli and ask the respondent for same/different judgments while recording both the correctness and the response time for each pair (e.g., Thomas, 1996). Another common type of application, often used in marketing or political science, has respondents make preference judgments (e.g., Richards, 1997). For example, respondents may be shown pairs of competing products or political candidates and asked to select the one that they prefer. In the measurement of attitudes, interests, or personality traits, paired comparisons are often called *forced choice items* (Gay, Weiss, Hendel, Dawis, & Lofquist, 1971; Jackson, 1977; Myers, 1995; Raskin &

M. L. Davison (✉) · A. Zieffler · J. Cabrera · S. R. Karl
Department of Educational Psychology, University of Minnesota,
Minneapolis, MN, USA
e-mail: mld@umn.edu

H. S. Cohen
AT&T Labs,
Middletown, NJ, USA

Hall, 1981). In forced choice items, respondents are provided two statements and asked to choose the statement in each pair with which they most agree or the statement that best describes them. In the fourth application of paired comparisons, respondents are shown two stimuli and asked to select the one perceived to be higher on a dimension specified in the questionnaire (e.g., Stylianou, 2003). For instance, the respondent might evaluate whether the crime of shoplifting or jaywalking is more “serious.”

For purposes of controlling order effects, Ross (1934) described two criteria for an optimal presentation of pairs: spacing and balancing. He described an algorithm for ordering all possible pairs, hereafter called the *optimally spaced ordering*. *Optimally spaced* means that appearances of a given stimulus are approximately equally spaced throughout the list of pairs. When n is odd and there is optimal spacing, the appearances of a stimulus will be separated by at least $(n - 3)/2$ pairs and by at most $(n - 1)/2$ pairs. If n is even, the appearances of a stimulus will be separated by at least $(n - 4)/2$ pairs and by at most $(n - 2)/2$ pairs. Ross also recommended balancing the pairs so that each stimulus appears equally often as the first or second member of a pair. He did not, however, propose an algorithm for balancing pairs. Cohen and Davison (1973) programmed algorithms for both the optimal spacing and the balancing of paired stimuli (although the balancing algorithm was not published).

To illustrate optimal spacing and balancing, Table 1 shows the balanced, optimally spaced ordering of pairs for nine university stimuli: 1 = *University of Chicago*, 2 = *University of Illinois*, 3 = *University of Michigan*, 4 = *University of Minnesota*, 5 = *Northwestern University*, 6 = *Purdue University*, 7 = *University of Wisconsin*, 8 = *Indiana University*, and 9 = *University of Iowa*. The pairs might be used in a study of institutional prestige in which an individual responds to each of the $n(n - 1)/2 = 36$ pairs by selecting the member that the individual considers the more prestigious.

The 36 optimally spaced, balanced pairs of stimuli appear from left to right across each row and then down the columns of Table 1. That is, the first four pairs appear from left to right in the first row, Pairs 5–8 appear from left to right in the second row, Pairs 9–12 appear from left to right in the third row, and so forth. Three characteristics of the pairs are of note. First, the pairs are optimally spaced, in that the appearances of a stimulus are separated by at least $(n - 3)/2 = 3$ pairs and at most $(n - 1)/2 = 4$ pairs. For instance, the appearances of the University of Chicago are always separated by three or four pairs, since it appears in the 1st, 5th, 10th, 14th, 19th, 23rd, 28th, and 32nd pairs. Second, the pairs are balanced, because each stimulus appears an equal number of times as the first (top) and the second (bottom) member of a pair. Note that

the University of Chicago appears in eight pairs, four times as the top member and four times as the bottom member. Third, the pairs display features not prominently discussed by Ross (1934), but which are quite useful in implementing matrix-sampling or incomplete designs: There are $n = 9$ blocks of pairs, each block contains $(n - 1)/2 = 4$ pairs, and each stimulus appears at most once in each block. In Table 1, the blocks correspond to rows and are labeled B_1 to B_9 . The University of Chicago appears once and only once in blocks B_1 to B_8 , but does not appear in B_9 .

Table 2 shows an optimally spaced, balanced ordering for an even number of stimuli, $n = 8$. It displays two of the three characteristics seen in Table 1. First, appearances of a stimulus are optimally spaced, in that they are separated by at least $(n - 4)/2 = 2$ pairs and at most $(n - 2)/2 = 3$ pairs. This can be seen in the appearances of the University of Chicago in the 1st, 4th, 8th, 11th, 15th, 18th, and 22nd pairs. Second, as nearly as possible, the pairs are balanced: Each stimulus appears almost equal numbers of times as the top and the bottom member of a pair. For instance, the University of Chicago appears as the top stimulus four times and as the bottom stimulus three times. When n is even, every stimulus appears in an odd number of pairs, $(n - 1)$ pairs, and therefore cannot appear as the top- and bottom-most member the exact same number of times. In a balanced ordering when n is even, the numbers of times that a stimulus appears first and second will differ by only one.

The pairs in Table 2 do not have the third feature seen in Table 1. Note, for instance, that the University of Chicago appears twice in the first block. Every stimulus, except the University of Illinois, appears twice in one of the blocks. Because the optimally spaced, balanced ordering for an even number of stimuli lacks this third property (every stimulus appears no more than once in a block), matrix-sampling and incomplete designs are more difficult to construct.

We now turn our attention to situations in which the number of stimuli is sufficiently large that it is unreasonable to ask respondents to judge all possible pairs. For such situations, we first describe the implementation of a matrix-sampling design in which the full list of pairs is divided into mutually exclusive, exhaustive subsets. Each person judges a sample of the pairs, and every pair is judged by a sample of people. Next, we turn to the implementation of fully incomplete designs, in which a subset of stimuli is chosen and only that subset is judged.

We then describe the algorithm for creating an optimally spaced, balanced ordering of pairs. The technical details of the algorithm can be skipped without loss of generality. The algorithm employs Ross's (1934) method for arranging pairs so that appearances of each stimulus are optimally spaced. To his method, we add an algorithm for balancing

Table 1 Optimally spaced and balanced ordering for all possible pairs of nine universities

Block	Paired Comparisons			
	1	2	3	4
B₁	University of Chicago University of Illinois	University of Iowa University of Michigan	Indiana University University of Minnesota	University of Wisconsin Northwestern University
B₂	Purdue University University of Chicago	University of Michigan University of Illinois	University of Minnesota University of Iowa	Northwestern University Indiana University
B₃	Purdue University University of Wisconsin	University of Chicago University of Michigan	University of Illinois University of Minnesota	University of Iowa Northwestern University
B₄	Indiana University Purdue University	University of Wisconsin University of Chicago	University Minnesota University of Michigan	Northwestern University University of Illinois
B₅	Purdue University University of Iowa	University of Wisconsin Indiana University	University of Chicago University of Minnesota	University of Michigan Northwestern University
B₆	University of Illinois Purdue University	University of Iowa University of Wisconsin	Indiana University University of Chicago	Northwestern University University of Minnesota
B₇	Purdue University University of Michigan	University of Wisconsin University of Illinois	Indiana University University of Iowa	University of Chicago Northwestern University
B₈	University of Minnesota Purdue University	University of Michigan University of Wisconsin	University of Illinois Indiana University	University of Iowa University of Chicago
B₉	Northwestern University Purdue University	University of Minnesota University of Wisconsin	University of Michigan Indiana University	University of Illinois University of Iowa

the pairs. The balancing algorithm is based on an unpublished method first programmed by Cohen and Davison (1973). Finally, a small modification of the algorithm is proposed, a modification that facilitates implementation of matrix-sampling and incomplete designs when n is even.

Matrix sampling

In some situations, the number of paired comparisons is too large to be feasibly judged by any one person. In such cases, the researcher may want to adopt a matrix-sampling design in which the questionnaire has several forms, with each form containing a different subset of the pairs. Each form would be administered to a different random sample of respondents. At its most basic level, our approach to constructing forms is rather simple. Instead of operating directly on individual pairs of stimuli and dividing the pairs evenly across forms, our approach utilizes the natural blocks of pairs in the optimally spaced, balanced ordering

and divides the blocks evenly across forms. For instance, if there are six blocks and two forms, the forms can be created by assigning the pairs of stimuli in **B₁**, **B₂**, and **B₃** to Form 1 and the pairs of stimuli in **B₄**, **B₅**, and **B₆** to Form 2. Three forms could also be created by assigning the pairs of stimuli in **B₁** and **B₂** to Form 1, **B₃** and **B₄** to Form 2, and **B₅** and **B₆** to Form 3. In some cases, however, the blocks do not distribute evenly across forms (e.g., two forms and seven blocks). In such cases, one can distribute all but the last block evenly over forms and then divide the pairs of stimuli in the last block over forms. For instance, with two forms and seven blocks, one could assign the pairs of stimuli in **B₁**, **B₂**, and **B₃** to Form 1, and the pairs of stimuli in **B₄**, **B₅**, and **B₆** to Form 2, and then evenly divide the pairs of stimuli in **B₇** across Forms 1 and 2.

By assigning blocks of pairs from an optimally spaced and balanced ordering, one can conveniently create forms that satisfy, as closely as possible, the following desirable conditions:

- Each pair of stimuli appears on exactly one form;
- Every form contains the same number of pairs;

Table 2 Optimally spaced and balanced ordering for all possible pairs of eight universities (excluding the University of Iowa)

Block	Paired Comparisons			
	1	2	3	4
B₁	University of Chicago University of Illinois	Indiana University University of Minnesota	University of Wisconsin Northwestern University	Purdue University University of Chicago
B₂	University of Michigan University of Illinois	Northwestern University Indiana University	Purdue University University of Wisconsin	University of Chicago University of Michigan
B₃	University of Illinois University of Minnesota	Indiana University Purdue University	University of Wisconsin University of Chicago	University of Minnesota University of Michigan
B₄	Northwestern University University of Illinois	University of Wisconsin Indiana University	University of Chicago University of Minnesota	University of Michigan Northwestern University
B₅	University of Illinois Purdue University	Indiana University University of Chicago	Northwestern University University of Minnesota	Purdue University University of Michigan
B₆	University of Wisconsin University of Illinois	University of Chicago Northwestern University	University of Minnesota Purdue University	University of Michigan University of Wisconsin
B₇	University of Illinois Indiana University	Northwestern University Purdue University	University of Minnesota University of Wisconsin	University of Michigan Indiana University

- Every stimulus appears an equal number of times on each form;
- The appearances of a stimulus are evenly spaced throughout the form; and,
- Across all forms (although not within a single form) each stimulus will appear equally as often as the first and as the second member of a pair.

Odd number of stimuli

We begin by considering the situation in which the number of stimuli is odd. To design a matrix sampling for an odd number of stimuli, one must first choose the number of forms f . Then one must evenly divide the stimuli in the last block across the forms and then evenly divide the remaining blocks across the forms.

As discussed above, when n is odd, the ordering will contain n blocks of consecutive pairs—each block containing $(n - 1)/2$ pairs. In Table 1, each block of four pairs corresponds to a row, with **B_i** designating the block in row i . Each block contains every stimulus except one, and the missing stimulus is different for each block. For instance, in Table 1, every stimulus except Purdue University appears once in **B₁**, every stimulus except the University of Wisconsin appears once in **B₂**, and so forth. Because the blocks have this structure, forms satisfying the criteria above can be assembled by a judicious assignment of

blocks to forms and by dividing the stimuli in the last block across forms. To obtain the desirable features described above, however, one must first carefully select the number of forms f so that the first $(n - 1)$ blocks and the stimuli in the last block, **B_n**, will divide evenly across forms.

Choosing the number of forms If n is an odd number, f forms can be created by choosing f such that it is a factor of $(n - 1)/2$. Such a choice of f is ideal for our purposes because both the first $(n - 1)$ blocks and the $(n - 1)/2$ pairs in the last block can then be divided evenly over the f forms. As a result, every pair will appear on exactly one form, an equal number of pairs will appear on each form, and an equal number of appearances of each stimulus will appear on each form. There will always exist at least one such factor, $f = (n - 1)/2$, since $(n - 1)$ is even if n is odd and $(n - 1)/2$ is a factor of itself. Each of the f forms will contain $(n - 1)/f$ blocks of the first $(n - 1)$ blocks and $(1/2)(n - 1)/f$ of the pairs in the last block, **B_n**.

Assigning blocks to forms In order for each stimulus to appear an equal number of times on each form, it is best to begin by distributing the pairs in the last block evenly across forms. When n is odd, each block is missing one stimulus, in the sense that there is one stimulus that does not appear in any pair of that block. For purposes of describing the assignment of blocks to forms when n is odd, it is useful to

subscript blocks somewhat differently than elsewhere in this article. Let $\mathbf{B}_i$ ($i \in 1, 2, \dots, n$) be the block corresponding to stimulus s_i , in the sense that stimulus s_i does not appear as a member of any pair in $\mathbf{B}_i$. After assigning pairs from the last block to a form, let s_i ($i \in 1, 2, \dots, n$) be one of the stimuli that appears in a pair from the last block, now assigned to a particular form. The remainder of the form should be composed of blocks $\mathbf{B}_i$ corresponding to stimuli s_i . For instance, if the pairs (s_1, s_8) and (s_2, s_7) in the last block have been assigned to Form 1, these pairs contain the stimuli s_1 , s_2 , s_7 , and s_8 , and consequently the pairs in blocks $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_7$, and $\mathbf{B}_8$ will be assigned to Form 1. That is, the blocks missing s_1 , s_2 , s_7 , and s_8 should be assigned to Form 1. Similarly, if the pairs (s_3, s_6) and (s_4, s_5) from the last block have been assigned to Form 2, these pairs contain the stimuli s_3 , s_4 , s_5 , and s_6 , and consequently the pairs in blocks $\mathbf{B}_3$, $\mathbf{B}_4$, $\mathbf{B}_5$, and $\mathbf{B}_6$ should be assigned to Form 2. That is, the blocks missing s_3 , s_4 , s_5 , and s_6 should be assigned to Form 2. The remaining forms are created in exactly the same way, based on the pairs in the last block assigned to the form.

Example For purposes of illustration, assume that we wish to create matrix-sampling forms for the set of nine stimuli in Table 1. With nine stimuli, creating either $f = 2$ or 4 forms is ideal; that is, both of these numbers are factors of $(n - 1)/2 = 4$. For purposes of this illustration, we choose $f = 2$. Hence, we will construct two forms, each having half of the stimulus pairs. In assigning pairs from the last block to forms, the researcher may want to divide the pairs in the last block evenly across forms using random assignment, but random assignment is not necessary to achieve forms with the properties described above. In this example, we have simply assigned the first half of the pairs from the last block to Form 1 and the last half of the pairs from the last block to Form 2.

In Step 1, we would divide the pairs in the last block $\mathbf{B}_9$ evenly across Forms 1 and 2. Let's assign the first two pairs of $\mathbf{B}_9$ (Northwestern University/Purdue University and University of Minnesota/University of Wisconsin) to Form 1 and the last two pairs (University of Michigan/Indiana University and University of Illinois/University of Iowa) to Form 2. This assignment of pairs now determines which of the remaining blocks $\mathbf{B}_1, \dots, \mathbf{B}_8$ will be assigned to Forms 1 and 2. Of the four pairs of stimuli in $\mathbf{B}_9$ assigned to Form 1 (Northwestern University, Purdue University, University of Minnesota, and University of Wisconsin), Northwestern University is missing from $\mathbf{B}_8$, Purdue University is missing from $\mathbf{B}_1$, University of Minnesota is missing from $\mathbf{B}_7$, and University of Wisconsin is missing from $\mathbf{B}_2$; hence, the pairs in blocks $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_7$, and $\mathbf{B}_8$ will be included in Form 1. Similarly, of the four pairs of stimuli in $\mathbf{B}_9$ assigned to Form 2 (University of Michigan, Indiana University, University of Illinois, and University of Iowa), University

of Michigan is missing from the pairs in $\mathbf{B}_6$, Indiana University is missing from the pairs in $\mathbf{B}_3$, University of Illinois is missing from $\mathbf{B}_5$, and University of Iowa is missing from the pairs in $\mathbf{B}_4$; hence, the pairs in blocks $\mathbf{B}_3$, $\mathbf{B}_4$, $\mathbf{B}_5$, and $\mathbf{B}_6$ will be included in Form 2.

Table 3 shows the pairs included on the two forms. The first form consists of $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_7$, $\mathbf{B}_8$, and the first two pairs of stimuli from $\mathbf{B}_9$. In a similar fashion, the second form consists of $\mathbf{B}_3$, $\mathbf{B}_4$, $\mathbf{B}_5$, $\mathbf{B}_6$, and the last two pairs of stimuli from $\mathbf{B}_9$. Every pair appears on exactly one form. Furthermore, both forms have equal numbers of pairs (18), and each stimulus appears exactly four times on each form. If one thinks of the two pairs from $\mathbf{B}_9$ as a miniblock at the end of the form, then the appearances of stimuli are spaced such that every stimulus appears exactly once in all blocks except one. Across the two forms (but not within forms), the pairs are also balanced so that every stimulus is presented four times as the first stimulus and four times as the second stimulus.

Even number of stimuli

If the number of stimuli n is even, a complexity arises. The $n(n - 1)/2$ pairs in the optimally spaced, balanced ordering can be divided into $(n - 1)$ blocks of $(n/2)$ pairs, but the stimuli are not ideally distributed across blocks. Ideally, a stimulus would appear once and only once in each block. For purposes of creating matrix-sampling forms, the ordering of the pairs must be reblocked so that every stimulus appears exactly once in each block. This can be accomplished fairly simply by moving the first pair in the ordering to the end of the ordering. That is, Pair 1 then becomes Pair $(n - 1)(n/2)$, and all other pairs move up one position in the ordering. Table 4 shows the optimal blocking for the eight stimuli in Table 2. Note that the first pair in Table 2 is now the last pair in Table 4, and all other pairs have been moved up one position. Also, note that each stimulus appears once and only once in each block. Hereafter, if stimuli are blocked so each stimulus appears once and only once in a block, the pairs are said to be optimally reblocked.

In creating forms using matrix sampling for an even number of stimuli, two cases must be considered, the case in which $(n - 1)$ is *not* a prime number and the case in which $(n - 1)$ is a prime number.

When $(n - 1)$ is not prime Since there are $(n - 1)$ blocks if n is even, the blocks can be distributed evenly over f forms if f is chosen such that $(n - 1)$ is evenly divisible by f . Since each stimulus appears only once in each block (after optimally reblocking as illustrated in Table 4) and each block has the same number of pairs, distributing blocks evenly over forms also distributes pairs of stimuli evenly over forms. For example, if $n = 10$, then $(n - 1) = 9$, and $f =$

Table 3 Two matrix-sampling forms constructed from the optimally spaced and balanced ordering of nine universities in Table 1

Pairs Assigned to Form 1				
B₁	University of Chicago University of Illinois	University of Iowa University of Michigan	Indiana University University of Minnesota	University of Wisconsin Northwestern University
B₂	Purdue University University of Chicago	University of Michigan University of Illinois	University of Minnesota University of Iowa	Northwestern University Indiana University
B₇	Purdue University University of Michigan	University of Wisconsin University of Illinois	Indiana University University of Iowa	University of Chicago Northwestern University
B₈	University of Minnesota Purdue University	University of Michigan University of Wisconsin	University of Illinois Indiana University	University of Iowa University of Chicago
B₉	Northwestern University Purdue University	University of Minnesota University of Wisconsin		
Pairs Assigned to Form 2				
B₃	Purdue University University of Wisconsin	University of Chicago University of Michigan	University of Illinois University of Minnesota	University of Iowa Northwestern University
B₄	Indiana University Purdue University	University of Wisconsin University of Chicago	University of Minnesota University of Michigan	Northwestern University University of Illinois
B₅	Purdue University University of Iowa	University of Wisconsin Indiana University	University of Chicago University of Minnesota	University of Michigan Northwestern University
B₆	University of Illinois Purdue University	University of Iowa University of Wisconsin	Indiana University University of Chicago	Northwestern University University of Minnesota
B₉	University of Michigan Indiana University	University of Illinois University of Iowa		

3 divides evenly into $(n - 1)$. One could then create three forms, each containing $(n - 1)/f = 3$ blocks of pairs. That is, using the optimally reblocked ordering of paired stimuli, Form 1 might contain the pairs in blocks **B₁**, **B₂**, and **B₃**; Form 2 might contain the pairs in **B₄**, **B₅**, and **B₆**; and Form 3 might contain the pairs in **B₇**, **B₈**, and **B₉**. The resulting forms would be as optimally balanced as possible, in that every pair would appear on exactly one form, every form would have the same length (15 pairs), and every stimulus would appear exactly three times on every form.

When $(n - 1)$ is prime In this case, the number of forms f can be chosen such that f is a factor of $(n - 2)$ and $(n/2)$, in which case, the first $(n - 2)$ blocks and the $(n/2)$ pairs in the last block can be evenly divided over the forms. The forms will then have an equal number of pairs and a nearly equal number of appearances of each stimulus on each form. In the first step, the pairs from the last block are distributed evenly across the forms. Then, in the second

step, the remaining $(n - 2)$ blocks are distributed evenly over the forms.

Example Consider the example from Table 4, in which $n = 8$. Since $(n - 1)$ is a prime number, f can be chosen as $f = 2$, since $f = 2$ is a factor of both $(n - 2) = 6$ and $(n/2) = 4$. In Step 1, the four pairs in the last block, **B₇**, would be distributed evenly over the two forms. In Step 2, the remaining six blocks would then be evenly distributed over the two forms. That is, using the example in Table 4, Form 1 could contain the pairs from blocks **B₁**, **B₂**, and **B₃**, plus the first two pairs in **B₇**. Form 2 would then contain the pairs from blocks **B₄**, **B₅**, and **B₆**, plus the last two pairs in **B₇**. Every pair would appear on exactly one of the two forms. Both forms would have the same number of pairs (14). Each stimulus would not appear the same number of times on each form, since each stimulus appears in $(n - 1) = 7$ pairs, and seven is not evenly divided across two forms. The number of appearances for any stimulus will differ across

Table 4 Optimally reblocked arrangement for all possible pairs of eight universities (excluding the University of Iowa), in which each stimulus appears once and only once in each block

Block	Paired Comparisons			
	1	2	3	4
B₁	Indiana University University of Minnesota	University of Wisconsin Northwestern University	Purdue University University of Chicago	University of Michigan University of Illinois
B₂	Northwestern University Indiana University	Purdue University University of Wisconsin	University of Chicago University of Michigan	University of Illinois University of Minnesota
B₃	Indiana University Purdue University	University of Wisconsin University of Chicago	University of Minnesota University of Michigan	Northwestern University University of Illinois
B₄	University of Wisconsin Indiana University	University of Chicago University of Minnesota	University of Michigan Northwestern University	University of Illinois Purdue University
B₅	Indiana University University of Chicago	Northwestern University University of Minnesota	Purdue University University of Michigan	University of Wisconsin University of Illinois
B₆	University of Chicago Northwestern University	University of Minnesota Purdue University	University of Michigan University of Wisconsin	University of Illinois Indiana University
B₇	Northwestern University Purdue University	University of Minnesota University of Wisconsin	University of Michigan Indiana University	University of Chicago University of Illinois

forms, but for any pair of forms, the number of appearances of a given stimulus will differ by at most one appearance. In this example, each stimulus would appear four times on one of the forms and three times on the other. Across all forms, the pairs will be as nearly balanced as possible.

In summary, if n is odd, we recommend choosing the number of forms f so that it is a factor of $(n-1)/2$. Then the f forms can be created by first distributing the pairs in the last block evenly over the forms and then distributing the remaining blocks evenly over forms, as described above. If n is even, the pairs of stimuli must first be optimally reblocked. If n is even and $(n-1)$ is not a prime number, one can first choose the number of forms f to be a factor of $(n-1)$ and then distribute the blocks evenly over the forms. If n is even and $(n-1)$ is a prime number, we recommend choosing f to be a factor of $(n-2)$ and $(n/2)$, distributing the stimuli in the last block evenly over the forms, and then distributing the remaining blocks evenly over the forms as described above.

Fully incomplete design

In the matrix-sampling design described above, all possible stimulus pairs are judged, but not by the same people. In a fully incomplete design, not all stimulus pairs are judged. In cases where n is large and the complete list of pairs is too

long to be judged by any one person, a fully incomplete design can be utilized in which every respondent judges the same subset of all possible pairs. This might be the case if the researcher were interested in individual differences and did not want spurious individual differences to be induced by variation in the pairs judged by the various individuals. In the subset of pairs judged, every stimulus should appear at least once and, ideally, every stimulus would also appear an equal number of times. These criteria can be completely met when n is odd by including the pairs on one or more forms from a matrix-sampling design. When n is even, the criteria can be met by including the stimuli from a subset of the blocks after optimally reblocking as described above.

If n is odd, sampling blocks will not produce the desired result, because each block is missing one stimulus. If, however, the researcher creates forms as described in the matrix-sampling design and then assembles the pairs from one or more of those forms for inclusion on the final questionnaire, the resulting questionnaire will have the desirable feature that all stimuli will appear an equal number of times. For instance, if all respondents were administered Form 1 (or Form 2) in Table 3, each respondent would judge the same 18 pairs of stimuli. Every stimulus would appear on the questionnaire four times, and the appearances of each stimulus would also be evenly spaced over the questionnaire, in the sense that it would appear exactly once in all but one block (counting the miniblock at the end).

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 4 & 1 & \dots & \frac{n+1}{2} \\ n & n & 2 & 2 & 3 & 3 & 4 & \dots & \frac{n-1}{2} \\ n-1 & n-1 & n & n & 2 & 2 & 3 & \dots & \frac{n-3}{2} \\ n-2 & n-2 & n-1 & n-1 & n & n & 2 & \dots & \frac{n-5}{2} \\ n-3 & n-3 & n-2 & n-2 & n-1 & n-1 & n & \dots & \frac{n-7}{2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \frac{n+3}{2} & 1 & \frac{n+5}{2} & 1 & \frac{n+7}{2} & 1 & \dots & n \end{bmatrix}$$

Fig. 1 Matrix $\mathbf{X}$, containing the first element in each pair of an optimally spaced ordering for n stimuli

If n is even, then after optimally reblocking as illustrated in Table 4, sampling two or more blocks for inclusion on the incomplete questionnaire will accomplish the desired purpose. For instance, consider the situation where $n = 20$, which results in 190 pairs of stimuli. The researcher may decide that 100 paired comparisons are as many as any one respondent can be expected to judge. In the optimal reblocking, there will be $(n - 1) = 19$ blocks, each containing $(n/2) = 10$ pairs of stimuli. By sampling 10 of the 19 blocks, the researcher can create a questionnaire with 100 pairs in which each stimulus appears exactly 10 times. The appearances of each stimulus will be evenly spaced, in the sense that each stimulus will appear once and only once in each block.

Ordering and balancing a complete list of pairs

All of the procedures described above for a complete design, a matrix-sampling design, or a fully incomplete design presume the existence of an optimally spaced, balanced ordering of pairs of stimuli. In this section, we begin by describing Ross's (1934) algorithm for spacing the pairs. Then we describe the unpublished algorithm of Cohen and Davison (1973) for balancing the pairs. Those less interested in the technical details of the algorithms may skip to the Conclusions section without loss of continuity.

In what follows, we only describe the algorithm for optimally spacing n , an odd number of stimuli. If the number of stimuli is even, then a dummy stimulus is added to make an odd number of stimuli, the optimally spaced ordering is obtained for $(n + 1)$ stimuli (including the dummy stimulus), and after balancing the pairs, all pairs involving the dummy stimulus are removed. In what follows, each stimulus is designated a numeric code indicating its ordinal position in the list of stimuli.

For n stimuli, optimal spacing begins with the construction of two $[(n + 1)/2] \times (n - 1)$ matrices, $\mathbf{X}$ and $\mathbf{Y}$. Once constructed, the elements of $\mathbf{X}$ will correspond to the first

stimulus that should be presented for each pair, and the elements of $\mathbf{Y}$ will correspond to the second stimulus presented for each pair.

Matrix $\mathbf{X}$

The pattern to create matrix $\mathbf{X}$ is shown in Fig. 1; let r and c be subscripts designating a row or a column of $\mathbf{X}$, respectively.

The first row of $\mathbf{X}$ is a sequence of number pairs with a 1 in each odd-numbered column and consecutive integers from 2 through $(n + 1)/2$ in the even-numbered columns. The last row simply continues the sequence in the first row. That is, the first and last rows can be described as follows: If $r = 1$ or $r = (n + 1)/2$, then for any value of c ,

$$X_{r,c} = \begin{cases} 1 & \text{if } c \text{ is odd,} \\ r + \left(\frac{1}{2}\right)c & \text{if } c \text{ is even.} \end{cases} \quad (1)$$

For the remaining rows, $1 < r < (n + 1)/2$, the first two elements for each row are computed using,

$$X_{r,c} = n - r + 2. \quad (2)$$

The remaining elements in each row are then consecutively increasing integers, each repeated twice, continuing up to n , and then starting over with 2 and ending at the end of each row. For example, when $n = 9$, the third row would start with $(n - r + 2) = 8$ repeated twice, followed by 9 repeated twice, followed by 2 repeated twice, and 3 repeated twice; $\mathbf{X}$ is shown below for this example. (The triangular brackets used in the last row of matrices $\mathbf{X}$ and $\mathbf{Y}$ are explained below.)

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 4 & 1 & 5 \\ 9 & 9 & 2 & 2 & 3 & 3 & 4 & 4 \\ 8 & 8 & 9 & 9 & 2 & 2 & 3 & 3 \\ 7 & 7 & 8 & 8 & 9 & 9 & 2 & 2 \\ 1 \langle 6 \rangle 1 \langle 7 \rangle 1 \langle 8 \rangle 1 \langle 9 \rangle \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} 2 & 3 & 3 & 4 & 4 & 5 & 5 & \dots & \frac{n+3}{2} \\ 3 & 4 & 4 & 5 & 5 & 6 & 6 & \dots & \frac{n+5}{2} \\ 4 & 5 & 5 & 6 & 6 & 7 & 7 & \dots & \frac{n+7}{2} \\ 5 & 6 & 6 & 7 & 7 & 8 & 8 & \dots & \frac{n+9}{2} \\ 6 & 7 & 7 & 8 & 8 & 9 & 9 & \dots & \frac{n+11}{2} \\ \frac{n+3}{2} & \frac{n+5}{2} & \frac{n+7}{2} & \frac{n+9}{2} & \frac{n+11}{2} & \frac{n+13}{2} & \frac{n+15}{2} & \dots & n+1 \end{bmatrix}$$

Fig. 2 Matrix $\mathbf{Y}$, containing the second element in each pair of an optimally spaced ordering for n stimuli

Matrix $\mathbf{Y}$

For n stimuli, Fig. 2 shows the pattern used to create matrix $\mathbf{Y}$. The first element of each row in $\mathbf{Y}$ is computed using the equation

$$Y_{r,1} = r + 1. \quad (3)$$

The remaining elements in each row are then consecutively increasing integers, each repeated twice, except the last column. For instance, the first row starts with $(r + 1) = 2$, followed by 3 repeated twice, 4 repeated twice, 5 repeated twice, and then 6. Similarly, the third row begins with $(r + 1) = 4$, followed by 5 repeated twice, 6 repeated twice, 7 repeated twice, and then 8. $\mathbf{Y}$ is shown below for the situation in which $n = 9$.

$$\mathbf{Y} = \begin{bmatrix} 2 & 3 & 3 & 4 & 4 & 5 & 5 & 6 \\ 3 & 4 & 4 & 5 & 5 & 6 & 6 & 7 \\ 4 & 5 & 5 & 6 & 6 & 7 & 7 & 8 \\ 5 & 6 & 6 & 7 & 7 & 8 & 8 & 9 \\ 6 & \langle 7 \rangle & 7 & \langle 8 \rangle & 8 & \langle 9 \rangle & 9 & \langle 10 \rangle \end{bmatrix}$$

Assembling the optimally ordered pairs

Each pair of stimuli will consist of an element from $\mathbf{X}$ and the corresponding element from $\mathbf{Y}$, such that the pair will be $(\mathbf{X}_{r,c}, \mathbf{Y}_{r,c})$. The pairs are composed beginning with $(\mathbf{X}_{1,1}, \mathbf{Y}_{1,1})$ and then continuing down to the end of the first column, and then repeating this process for each subsequent column. The pairs created using the elements in the even-numbered columns of the last row, however, are omitted (the omitted elements are depicted in the matrices above by triangular brackets). At this point, the pairs of stimuli are optimally spaced, but not yet balanced.

Table 5 shows the optimally spaced, unbalanced pairs for nine stimuli. The first column provides the pair number; the second column gives the element location in $\mathbf{X}$ and $\mathbf{Y}$; and the third and fourth columns offer the numeric codes for the two stimuli to be presented (again, note that these pairs are unbalanced).

Balancing

The following algorithm will balance the optimally spaced pairs so that each stimulus will appear equally often as the first and the second member of a pair for an odd number of stimuli.

The order of the stimuli presented is reversed starting with pair $(n + 1)/2$, in groups of $(n + 1)/2$ consecutive pairs. This is repeated after skipping $(n - 1)/2$ pairs in between

each reversed set. The total number of groups with reversed pairs is $(n - 3)/2$. In other words, the algorithm maintains the order for the first $(n - 1)/2$ pairs, then reverses the order of the next $(n + 1)/2$ pairs, maintains the order for the next $(n - 1)/2$ pairs, then reverses the order of the next $(n + 1)/2$ pairs, and so on. Additionally, if either stimulus in pair number $(n - 1)^2/2 = 1$, then the order of that particular pair will also be reversed (if neither stimulus is 1, the pair maintains its original order).

The fifth column of Table 5 shows which pairs were reverse-ordered and which were not. The sixth and seventh columns of the same table show the pairs after reversing the order within pairs. The pairs are now balanced and optimally ordered. The reader can readily verify (using columns 6 and 7) that each stimulus appears four times as the first member of a pair and four times as the second member.

Conclusions

According to Ross (1934), optimally arranging paired comparisons involves balancing stimuli within pairs and spacing stimuli appearances across pairs. Here, we have described Ross's algorithm for spacing stimuli appearances across pairs, as well as Cohen and Davison's (1973) unpublished algorithm for balancing pairs. The latter algorithm makes it possible to balance the pairs for any number of stimuli. We have also described an algorithm for optimally blocking pairs of stimuli when n is even to facilitate matrix-sampling designs and fully incomplete designs.

The optimally spaced ordering for a set of n stimuli is not unique. There will be $n!$ optimally spaced arrangements, one for each ordering of the n stimuli. Randomly permuting the order of the n stimuli and then deriving the optimally spaced arrangement for that stimulus ordering is tantamount to randomly selecting one of the $n!$ optimally spaced arrangements. If a researcher wishes to create several complete forms, each with a different optimally spaced ordering of the pairs, the researcher can randomly permute the ordering of the stimuli several times, derive the optimally spaced arrangement for each stimulus ordering, and thereby create several forms each with a different, randomly chosen optimally spaced ordering.

Starting from an optimally spaced, balanced ordering of all possible pairs, we have described how to assemble forms through matrix sampling so that every pair appears on exactly one form; there is an equal number of pairs on each form; as nearly as possible, each stimulus appears equally often on each form; appearances of a stimulus are spaced within a form; and the pairs are balanced across forms. By assembling the pairs from one or more forms (if n is odd) or one or more blocks (if n is

Table 5 Maximally spaced orderings of pairs, both unbalanced and balanced, for nine stimuli

Pair Number	Element in X and Y	Maximally Spaced Pairs		Pattern of Block Balancing	Balanced Pairs	
		Stimulus 1	Stimulus 2		Stimulus 1	Stimulus 2
1	[1,1]	1	2	Skip 4 rows, no reverse	1	2
2	[1,2]	9	3		9	3
3	[1,3]	8	4		8	4
4	[1,4]	7	5		7	5
5	[1,5]	1	6	Block 1 of 5 pairs, reverse	6	1
6	[2,1]	2	3		3	2
7	[2,2]	9	4		4	9
8	[2,3]	8	5		5	8
9	[2,4]	7	6		6	7
10	[3,1]	1	3	Skip 4 rows, no reverse	1	3
11	[3,2]	2	4		2	4
12	[3,3]	9	5		9	5
13	[3,4]	8	6		8	6
14	[3,5]	1	7	Block 2 of 5 pairs, reverse	7	1
15	[4,1]	3	4		4	3
16	[4,2]	2	5		5	2
17	[4,3]	9	6		6	9
18	[4,4]	8	7		7	8
19	[5,1]	1	4	Skip 4 rows, no reverse	1	4
20	[5,2]	3	5		3	5
21	[5,3]	2	6		2	6
22	[5,4]	9	7		9	7
23	[5,5]	1	8	Block 3 of 5 pairs, reverse	8	1
24	[6,1]	4	5		5	4
25	[6,2]	3	6		6	3
26	[6,3]	2	7		7	2
27	[6,4]	9	8		8	9
28	[7,1]	1	5	Skip 4 rows, no reverse	1	5
29	[7,2]	4	6		4	6
30	[7,3]	3	7		3	7
31	[7,4]	2	8		2	8
32	[7,5]	1	9	Row includes 1, reverse	9	1
33	[8,1]	5	6	Skip 4 rows, no reverse	5	6
34	[8,2]	4	7		4	7
35	[8,3]	3	8		3	8
36	[8,4]	2	9		2	9

even), one can create a design in which each respondent is presented an optimal, fully incomplete list of pairs. All stimuli will appear equal numbers of times in the

incomplete list, appearances of every stimulus will be spaced throughout the list, and, as nearly as possible, pairs will be balanced.

A random arrangement of stimuli within pairs and across pairs is probably the major alternative in practice for controlling order effects. That is, one can generate all possible pairs and then randomly order the pairs and randomly order the stimuli (first or second member) within pairs. Unless one implements the randomization via a computer, it is a tedious and error-prone process. Furthermore, randomization is a rather untrustworthy method of controlling order effects. Within pairs, it will balance the appearances of a stimulus as the first or second member of a pair only within the limits of random sampling. Likewise, it will space the appearances of a stimulus evenly over the list of pairs only within the limits of random sampling. With computerization, the optimally spaced, balanced ordering described above can be accomplished just as readily as randomization, and with far tighter control of order effects.

If a researcher plans to use matrix sampling to create forms, the researcher may want to choose the number of forms f so that there are ideal factors, making it possible to optimally construct matrix forms. If n is odd, there will always exist at least one ideal factor $f = (n - 1)/2$. If n is even and divisible by 4, there will always be at least one ideal factor $f = 2$. It is our conjecture that researchers may often be interested in using $f = 2$ forms, particularly if the number of stimuli is not very large; $f = 2$ will always be an ideal factor if n is odd and $(n - 1)$ is divisible by 4, or if n is even and divisible by 4. Thus, researchers may want to choose the number of stimuli so that $(n - 1)$ or n is divisible by 4.

The decision as to whether to employ an incomplete design will depend on several factors. First, the ages and abilities of respondents must be taken into account. While college students may be able to judge a large number of stimuli, children may not be able to do so. In a study of letter confusions, the researcher may not want to present all possible pairs of the 26 letters in the alphabet to first graders. The researcher may decide that each child should judge only a subset of the 338 pairs so as not to exceed the child's limited attention span. Judging all possible pairs might also pose a problem for elderly respondents or respondents with disabilities. A study of confusions among Chinese writing characters would pose even greater challenges that might require an incomplete design. Second, the process of stimulus presentation can sometimes be lengthy, making a large number of judgments difficult if not impossible. For instance, tasting pairs of wines would require cleansing the palate before tasting each wine, which would tend to make judging a large number of pairs into a lengthy process for any one person. Third, judging a large number of stimuli can become quite fatiguing if the stimuli are complex. If the pairs involve complex descriptions of consumer products (e.g., cars), respondents might be limited in the number of stimuli that

they can reasonably judge without risking fatigue effects or waning attention. Finally, the paired-comparison task might be only one task in the study's larger data collection process. Political telephone surveys may ask about pairs of politicians, but may also include other measures of attitudes, opinions, demographics, and so forth. The researcher may want to limit the time devoted to the paired-comparison task for any one respondent using a fully incomplete design so as to leave sufficient time for other measures.

However useful fully incomplete designs may be, they may not be equally appropriate for all situations. They seem least suited to situations in which there are interactions between the two stimuli in a pair, such that the perceived location of a stimulus varies depending on the other stimulus with which it is paired (i.e., assimilation or contrast effects). We conjecture that the interactions may tend to balance out over the appearances of a given stimulus in a complete design or a matrix-sampling design, but that they may not do so with fully incomplete designs. Therefore, parameter estimates for a given stimulus estimated from a fully incomplete design might be affected by the other stimuli with which the stimulus is paired (and not paired) in a given design.

Models for paired comparisons differ in the number of free parameters to be estimated. For instance, McGuire and Davison (1991) described a fixed-effects model. Bockenholt (2001, 2004) and Maydeu-Olivares and Bockenholt (2005) described random-effects models that greatly relax the assumptions of a fixed-effects model but contain additional variances and covariances to be estimated. As the number of parameters to be estimated increases, presumably the amount of data needed, both the number of respondents and/or the number of pairs judged by each respondent, will increase. For matrix-sampling designs and fully incomplete designs, it is unclear how much missing data and what patterns of missing data can be tolerated without serious deterioration of parameter estimates and their standard errors. Some of the interesting questions involve the degree to which more pair judgments per respondent can compensate for a small number of respondents, and vice versa. If one is going to sample blocks of pairs, it is unclear how many blocks would be needed to adequately estimate the parameters. These questions must remain a matter for future research, and the answers may vary depending on the number and nature (fixed-effect vs. random-effect variances/covariances) of the parameters to be estimated in the model.

Unless n is very small, constructing all possible pairs, optimally spaced and balanced or not, can be tedious and error-prone. Therefore, the authors have created a package, Optimal Paired Comparisons (OPC), for the R software environment (R Development Core Team, 2011) to imple-

ment the algorithms to create the optimally spaced, balanced ordering for a set of stimuli. The OPC package is available from R-Forge (<http://opc.r-forge.r-project.org>).

References

- Bockenholt, U. (2001). Hierarchical modeling of paired comparison data. *Psychological Methods*, 6, 49–66. doi:[10.1037/1082-989X.6.1.49](https://doi.org/10.1037/1082-989X.6.1.49)
- Bockenholt, U. (2004). Comparative judgments as an alternative to ratings: Identifying the scale of origin. *Psychological Methods*, 9, 453–465. doi:[10.1037/1082-989X.9.4.453](https://doi.org/10.1037/1082-989X.9.4.453)
- Cohen, H. S., & Davison, M. L. (1973). Jiffy-scale: A FORTRAN IV program for generating Ross-ordered pair comparisons. *Behavioral Science*, 18, 76.
- Davison, M. L., & Jones, L. E. (1976). A similarity attraction model for predicting sociometric choice from perceived group structure. *Journal of Personality and Social Psychology*, 33, 601–612.
- Gay, E. G., Weiss, D. J., Hendel, D. D., Dawis, R. V., & Lofquist, L. H. (1971). *Manual for the Minnesota Importance Questionnaire*. Minneapolis, MN: Industrial Relations Center, University of Minnesota.
- Jackson, D. N. (1977). *Jackson Vocational Interest Survey manual*. London, ON: Research Psychologists Press.
- Maydeu-Olivares, A., & Bockenholt, U. (2005). Structural equation modeling of paired-comparison and ranking data. *Psychological Methods*, 10, 285–304. doi:[10.1037/1082-989X.10.3.285](https://doi.org/10.1037/1082-989X.10.3.285)
- McGuire, D. P., & Davison, M. L. (1991). Testing group differences in paired comparisons data. *Psychological Bulletin*, 110, 171–182. doi:[10.1037/0033-2909.110.1.171](https://doi.org/10.1037/0033-2909.110.1.171)
- Myers, I. B. (1995). *Gifts differing: Understanding personality type*. Mountain View, CA: Davies-Black.
- R Development Core Team. (2011). R: A language and environment for statistical computing [Software environment]. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, available at www.R-project.org
- Raskin, R., & Hall, C. S. (1981). The *Narcissistic Personality Inventory*: Alternate form reliability and further evidence of construct validity. *Journal of Personality Assessment*, 45, 159–162.
- Richards, J. E. (1997). Effects of attention on infants' preference for briefly exposed visual stimuli in the paired-comparison recognition-memory paradigm. *Developmental Psychology*, 33, 22–31.
- Ross, R. T. (1934). Optimum orders for the presentation of pairs in the method of paired comparisons. *Journal of Educational Psychology*, 25, 375–382.
- Schubsachs, A. P. W., & Davison, M. L. (1979). Individual differences in perceptions of occupations and occupational reinforcers. *Journal of Occupational Psychology*, 52, 299–310.
- Stylianou, S. (2003). Measuring crime seriousness: What have we learned and what else do we want to know? *Journal of Criminal Justice*, 31, 37–56.
- Thomas, R. D. (1996). Separability and independence of dimensions within the same-different judgment task. *Journal of Mathematical Psychology*, 40, 318–341. doi:[10.1006/jmps.1996.0032](https://doi.org/10.1006/jmps.1996.0032)
- Thurstone, L. L. (1927a). A law of comparative judgment. *Psychological Review*, 34, 273–286. doi:[10.1037/h0070288](https://doi.org/10.1037/h0070288)
- Thurstone, L. L. (1927b). Psychophysical analysis. *The American Journal of Psychology*, 38, 368–389.