PERMUTATION-BASED METHODS FOR EXAMINING CONFUSION DATA IN ESP EXPERIMENTS

By Stephanie Stahl

ABSTRACT: Permutation-based methods have received some attention in the parapsychological literature. However, their potential for analyzing confusion matrices associated with ESP experiments has been relatively untapped. This paper presents permutation tests applied to confusion data in parapsychology. Some tests examine agreement of two or more matrices with no regard to inherent internal patterns within the matrices, whereas others test for the presence of more sophisticated structures. Other possible methods for extracting structures from one or more confusion matrices—seriation, cluster analysis, and scaling—are also described in this paper. The range of possibilities for permutation-based analyses applied to confusion in psi is limited only by the quantitative analyst's imagination. The methods presented in this paper are demonstrated using data gleaned from the parapsychological literature.

Although much of the emphasis in ESP studies involves statistical analysis of the direct hits on the selected targets, the importance of examining the structure of the misses has also been recognized (Rhine, 1952). For example, a considerable body of research has focused on the displacement effect, which suggests that subject calls might precede or succeed the presented targets by one or more places (Pratt & Foster, 1950; Russell, 1943). A review of the displacement trend is provided by Milton (1987). Crandall (1989) examined the displacement trend, noting the dry spell of a few decades since it had initially been reported in parapsychological literature.

A second, but less substantial, volume of research under the umbrella of psi-missing is associated with the study of confusion structures (Cadoret, 1957; Cadoret & Pratt, 1950; Kelly, 1980; Kelly, Kanthamani, Child, & Young, 1975; Kennedy, 1979; Timm, 1969). As these authors have observed, a systematic tendency to confuse certain targets with other targets can also be indicative of the presence of ESP. Cadoret and Pratt (1950) provided a significant contribution with their development of the *consistent missing* (CM) theory, which reflects the propensity of subjects to exhibit systematic patterns in their incorrect identifications of the targets. These authors also proposed a χ^2 test for determining whether or not such systematic patterns were present in the data. They noted the necessity of developing a "method of evaluating the misses that was independent of the

number of hits" (Cadoret & Pratt, 1950, p. 245). In reference to the Fisk and Mitchell (1953) and Fisk and West (1957) clock card experiments designed to examine misses, Kennedy noted that "the CM analysis that evaluates incorrect calls in isolation from direct hits has not been applied to most of these data" (Kennedy, 1979, p. 116). Kelly et al. extended the CM paradigm by comparing the resulting confusion structures of visual recognition and ESP tasks for the same exceptional subject. Kelly et al. used *nonmetric multidimensional scaling* (MDS) to reveal similarities between the two confusion structures. In light of the substantive findings that can be uncovered from confusion data, we should not be surprised to find that future analyses in this area were strongly encouraged by Burdick and Kelly (1977, p. 101): "The confusions methodology opens up a rich set of possibilities for investigating the mechanisms of psi, and we hope it will be vigorously pursued in future work."

Despite the suggestion for further development of methods for studying ESP confusion data, research in this area remains rather scant. This dearth of literature in parapsychology is in marked contrast to experimental psychology, as this latter area has a literature base that is replete with studies emphasizing the analysis of confusion structures. Psychological pplications include visual recognition of letters and/or digits (Townsend, 971), visual recognition of textures (Cho, Yang, & Hallett, 2000), lipreading tasks (Manning & Shofner, 1991), auditory recognition tasks (Morgan, Chambers, & Morton, 1973), taste recognition (Hettinger, Gent, Marks, & Frank, 1999), odor discrimination (Kent, Youngentob, & Sheehe, 1995), and tactile recognition (Vega-Bermudez, Johnson, & Hsiao, 1991). This body of research examines the study of confusion in sensory perception. Clearly, research should be further extended in *extra*sensory perception.

At least two important issues are associated with the study of confusion matrices obtained from ESP experiments: (a) the implementation of tests for the presence of systematic patterning in a single confusion matrix or among multiple confusion matrices, and (b) the deployment of analytical methods to uncover confusion structures that are masked by noisy data. In this paper, permutation-based methods are presented to tackle both of the issues related to analyzing confusion data from ESP experiments. Specifically, straightforward permutation tests are proposed for measuring the level of concordance among multiple confusion matrices with the same set of targets as well as for evaluating the symmetry of a single confusion matrix. An important aspect of one of these tests is the incorporation of information about the internal structural properties of each confusion matrix when testing concordance. For uncovering structure within one or more confusion matrices, I propose the implementation of seriation methods, cluster analysis, and scaling.

The next section of this article introduces permutation tests, beginning with subsections to introduce terminology and provide comparisons to familiar tests in parapsychology. The subsequent section

describes and demonstrates permutation tests for testing concordance of two (or more) matrices and for testing symmetry in a confusion matrix. Prior to the concluding section, a brief section discusses other methods for extracting structure from ESP confusion data—seriation, cluster analysis, and scaling—to optimally order targets according to confusion.

Permutation Tests to Examine Confusion in Parapsychological Experiments: Background Literature and Terminology

Permutation-based methods for significance testing have been used since the pioneering work of Pittman (1937) and Fisher (1935). Of course, Fisher is also known for his work in parapsychology (1928). However, Edgington (1964, 1966, 1969) is generally credited with introducing permutation-based methodology to psychology. Permutation tests are one of the permutation-based methodologies and are not new to parapsychological studies, dating as far back as Pratt and Birge (1948) who used permutation tests to assess verbal material from mediums.

A couple of important distinctions need to be made. First, the difference between permutation tests and randomization tests should be clarified. A randomization test is a type of permutation test that incorporates randomization. "Randomization test" originally referred to random assignment of subjects to treatments, but contemporary researchers use the term to refer to experiments in which stimuli are randomly selected (Bradley, 1968; Edgington, 1995). Therefore, for general purposes, this article uses the broader term "permutation test" and encourages parapsychologists to apply randomization where appropriate. The second clarification to be made is between confusion matrices and proximity matrices. Proximity matrices often record "closeness" of objects or stimuli in terms of similarity or dissimilarity. Confusion matrices record similarity (or how easily targets are confused) in non-negative terms, where a larger off-diagonal entry reflects greater confusion between two stimuli. Hence, confusion matrices are a class of similarity matrices within the broader context of proximity matrices.

In addition to similarity, confusion matrices are characterized by asymmetry. The asymmetry refers to the tendency in real data to not always record exact reverse confusion. For example, a circle might be called one half of the time when a square was actually shown, yet a square might be called only one third of the time when a circle was actually shown. Visually, the upper right half of the matrix does not mirror the lower left half of the matrix about the main diagonal.

The usual representation of correct and incorrect responses for the n stimuli in psychological experiments is an $n \times n$ confusion matrix. Rows are labeled according to targets or stimuli, whereas columns are labeled according to calls or responses. The main diagonal of the matrix contains the correct responses for each stimulus, whereas the incorrect/confused

responses correspond to the off-diagonal elements. Hence, the diagonal elements are often irrelevant in the study of confusion. In some cases, the entries in the confusion matrix are normalized to represent proportions rather than raw numbers. Within the context of ESP experiments, main diagonal entries are hits and off-diagonal entries are misses.

Many permutation tests involve the reordering of a matrix such that both columns and rows are permuted similarly. That is, the data remain accurately recorded but reordered, and the trace entries remain the same but are reordered. This is the method used in the cases of examining structure of matrices, concordance of matrices, and symmetry about the diagonal of a matrix.

Finally, the demonstrations in this paper are *exact* tests performed by total enumeration of all possible permutations of the targets. However, computer speed and storage (i.e., RAM) generally determine the feasibility of performing total enumeration for permutation tests with large target sizes. For $n > 13$, permutation tests for confusion are usually *approximate* tests, which entail random or pseudorandom sampling of permutations from the set of all permutations for the test statistic being used. One reviewer for this article recommended that the number of permutations should be at least 50 divided by the probability of interest, e.g., 50,000 permutations for evaluating $p = .001$ because $50{,}000 = 50/.001$.

Examples of Familiar Tests in Relation to Permutation Testing

Intuitively, exact permutation tests simply evaluate an observed outcome with respect to all possible alternative outcomes. Such tests have been employed by parapsychologists, and those familiar tests are helpful in understanding how permutation tests can be used to analyze confusion in psi.

A popular statistical test used to evaluate association between two categorical variables is Fisher's Exact Test. In parapsychology, relatively recent usage of Fisher's Exact Test is seen in the research of cerebral hemisphere dominance and ESP performance in the autoganzfeld (Alexander & Broughton, 2001) as well as a more ghostly application (Maher, 1999). Fisher's test assumes random sampling, a directional hypothesis, independent observation, mutual exclusivity, and a dichotomous level of measurement. Although Fisher initially examined the 2×2 case, this test can be extended to $m \times n$ matrices, however unwieldy that extended exercise might be. Fisher's Exact Test is salient in that it demonstrates how data in a matrix can be altered to provide important information about the matrix as a whole. Permutation tests can require order-preserving or order-reversing transformations, such as converting similarity/dissimilarity matrices to dissimilarity/similarity matrices. Other common practices involve transformation by row-normalization or converting asymmetric matrices to

symmetric matrices. However, analysts are cautioned that not all permutation tests can tolerate such transformations.

Morris (1972) described a useful method for evaluating free-response material, the preferential matching exact test with respect to Stuart (1942). In particular, this test uses rankings of targets and compares the sum of rankings to the number of possible rank sums of lesser value. Naturally, the desirable circumstance is when the number of possible ranks is equal to the number of trials. In this test, the free-response material is considered and ranked as a whole, rather than atomisticly broken into smaller units of information to be matched against characteristics of the targets. Solfvin, Kelly, and Burdick (1978) extended and generalized the work of Morris. Burdick and Kelly (1977) categorized the preferential ranking and rating methods as one of the two main subclasses of holistic approaches to analyzing parapsychological data, the other subclass being forced-choice techniques. These methodologies are appropriate for analyzing correct matchings; that is, the objective is to attain high ranks for targets with their correct responses/protocols. An alternate judging scheme is to have judges rank each target with each response/protocol (Schlitz & Gruber, 1980), which produces an $n \times n$ asymmetric matrix more akin to a confusion matrix.

Utts (1993, p.77) described an exact test presented by Scott (1972, p. 87) that has been used rather extensively in the parapsychological literature. The data are first arranged as detailed in the previous subsection with "hits" recorded along the diagonal. The diagonal entries (i.e. the trace) are summed. The columns are then reordered in all $n!$ (where $n! = n(n-1)\cdots 2\cdot 1$) possible permutations, with rows remaining in their original ordering and each trace summed. The statistic, which requires a random presentation of targets, is the proportion of sums that are as good as or better than the sum for the original "correct" ordering. This statistic was used by Targ (1994) in a remote-viewing replication study. Moreover, this counting measure has been discussed with regard to data analysis from Princeton Engineering Anomalies Research (Dobyns, Dunne, Jahn, & Nelson, 1992; Hansen, Utts, & Markwick, 1992), particularly with regard to the relevance of the diagonal entries. Unlike this counting measure in which only columns are permuted and, hence, the diagonal changes, the permutation tests presented in this paper require that the permutation be applied to both columns and rows, and the diagonal entries are unaltered although possibly reordered.

In this article, permutation tests are utilized to examine off-diagonal entries in ESP confusion matrices. In parapsychological studies, the more familiar use of permutation tests is in verification of, or comparison to, correlation statistics for analysis of hits or direct correlations of variables, such as the comparison made by Kennedy, Kanthamani, and Palmer (1994, p. 365) of a permutation method (Edgington, 1995) to the Pearson correlation coefficient. Schmidt, Schneider, Binder, Bürkle, and Walach

(2001) performed a battery of tests to answer important questions concerning the methodology and analysis of *direct mental interaction between living systems* (DMILS) via examination of *electrodermal activity* (EDS), including *percentage influence score* (PIS), the paired *t* test, the Wilcoxon signed rank test, and a two-component model of the Wilcoxon signed rank test and randomized permutation analysis. Their permutation testing was originally designed for waveform comparison testing (Blair & Karniski, 1993) and adapted to study DMILS data by Radin, Machado, & Zangari (1998). The results of Schmidt et al. (2001) indicated that the Wilcoxon signed rank test was the easiest test to perform without distribution assumptions. However, they noted: "Full (100%) power can be reached by the permutation test, but this method is not so easily applied and tends to take much time with large session numbers" (p.79).

With the advent of increased computer power available to researchers and with improvements in permutation-based methodologies, the full-power tests become more plausible. Because the focus of this paper is on permutation tests to analyze confusion, the examples are relatively small and easily allow for total enumeration in testing. A useful algorithm to systematically generate all permutations of n targets is provided in the Appendix. Although total enumeration produces $n!$ permutations to be evaluated, not all exact permutation tests require so many permutations.

Permutation Tests for Evaluating Agreement and Symmetry in Confusion Matrices

Tests Based on Agreement of Internal Confusion Structures

Once again referring to the psychological literature, comparisons of confusion among the senses have been made, such as Loomis (1982, p. 46, Table 5), who reported correlation coefficients between visual and tactile confusion matrices. With the aid of MDS, comparison between ESP and visual perception of playing cards has been reported (Kelly et al., 1975) in the parapsychological literature. For the ESP and visual perception data, the playing cards are the targets and responses of the confusion matrices yielding 13×13 matrices for the perception of numbers (ace through king) and 4×4 matrices for the perception of suits (spade, club, heart, diamond). Concordance can be more thoroughly examined in the parapsychological literature by scrutinizing the internal structure of confusion matrices.

Hubert (1978, 1987) presents a number of indices that capture internal structural characteristics within two or more matrices. One such index, applied to two matrices, is a within-stimulus gradient among triads of targets. As the name implies, relationships of responses (traveling down the permutation) to each particular target are examined, effectively looking at a confusion gradient along the permutation. The triad of targets is taken from both matrices to effectively compare the confusion gradients in the

matrices to each other. The null hypothesis states that the two matrices exhibit no concordance with respect to the patterning of elements within their rows, that is, confusion in the matrices does not follow a similar pattern across the targets. Like all permutation tests introduced in this article, the resulting p value is one-tailed. The following pseudocode can be easily inserted into the algorithm of the Appendix to calculate the within-stimulus triad test statistic and distribution.

```
Stat = 0
for i = 1:n
for j = 1:n (i ≠ j)
for k = 1:n (i ≠ k, j ≠ k)

 if C1[i, j] > C1[i , k] and
            C2 [Targets(i), Targets(j)] > C2 [Targets(i), Targets(k)] then
  Stat = Stat + 1    {Increment statistic as per "greater than" consistency.}
 if C1[i, j] < C1[i , k] and
            C2 [Targets(i), Targets(j)] < C2 [Targets(i), Targets(k)] then
  Stat = Stat + 1    {Increment statistic as per "less than" consistency.}
 if C1[i, j] > C1[i, k] and
            C2 [Targets(i), Targets(j)] < C2 [Targets(i), Targets(k)] then
  Stat = Stat – 1    {Decrement statistic as per inconsistency.}
 if C1[i, j] < C1[i, k] and
            C2 [Targets(i), Targets(j)] > C2 [Targets(i), Targets(k)] then
  Stat = Stat – 1    {Decrement statistic as per inconsistency.}

end for k
end for j
end for i
return Stat
```

To demonstrate the within-stimulus triad test, consider the 4 × 4 matrix for visual and ESP suit data from Kelly et al. (1975). Because confusion data are usually presented with rows labeled by stimuli, the within-stimulus triad test is effectively a within-row triad test. In Table 1, the ESP matrix is permuted, and corresponding pairs of entries within rows of the visual matrix— $\mathbf{C}_{\text{VISUAL}}$,—and the ESP matrix—$\mathbf{C}_{\text{ESP}}$, $\mathbf{C}_{\text{ESP(1)}}$, or $\mathbf{C}_{\text{ESP(2)}}$—are compared. Permutations were arbitrarily chosen as (C—D—S—H) and (C—S—D—H) to establish the row/column orderings of $\mathbf{C}_{\text{ESP(1)}}$ and $\mathbf{C}_{\text{ESP(2)}}$, respectively. (Notice in Table 1 that the data are not altered in the permuted matrices, merely reordered.) Similar comparisons (greater than/less than) are given a positive sign whereas dissimilar comparisons are given a negative sign. The signs are tallied to yield an index for each matrix relative to the visual matrix. The index reveals whether or not entries of two matrices tend to follow similar

patterns. The two permutations in this example illustrate the mediocrity of the index value for the ESP matrix with the same ordering, C—S—H—D, as the visual matrix. Ideally, we would list all possible permutations of the ESP matrix. If we perform this operation on all 4! = 24 permutations of the ESP matrix, we find that the indices range from –10 to 10. Essentially, we use the index for the pair with the same ordering, -2, as a baseline for comparison with all the calculated indices. Of the 24 possible permutations, 19 had a within-triad stimulus index of –2 or larger, giving a one-tailed p value of 19/24 = 0.7917. Clearly, the original ESP matrix, $\mathbf{C}_{ESP}$, shows no significant concordance with the visual matrix, $\mathbf{C}_{VISUAL}$, relative to the permuted matrices.

TABLE 1
CALCULATING THE CONCORDANCE INDEX BETWEEN VISUAL SUIT DATA AND ESP DATA

Visual (C_{VISUAL})				ESP (C_{ESP})				ESP ($C_{ESP(1)}$)				ESP ($C_{ESP(2)}$)			
	C	S	H	D	C	S	H	D	C	D	S	H	C	S	D	H
C		113	29	67	C		175	154	100	C		100	175	154	C		175	100	154
S	51		23	53	S	213		111	99	D	138		183	143	S	213			111
H	71	39		212	H	163	195		96	S	213	99		111	D	138	183		143
D	60	67	53		D	138	183	143		H	163	96	195		H	163	195	96	
113 > 29					175 > 154			(+1)		100 < 175			(-1)		175 > 100			(+1)	
113 > 67					175 > 100			(+1)		100 < 154			(-1)		175 > 154			(+1)	
29 < 67					154 > 100			(-1)		175 > 154			(-1)		100 < 154			(+1)	
51 > 23					213 > 111			(+1)		138 < 183			(-1)		213 > 99			(+1)	
51 < 53					213 > 99			(-1)		138 < 143			(+1)		213 > 111			(-1)	
23 < 53					111 > 99			(-1)		183 > 143			(-1)		99 < 111			(+1)	
71 > 39					163 < 195			(-1)		213 > 99			(+1)		138 < 183			(-1)	
71 < 212					163 > 96			(-1)		213 > 111			(-1)		138 < 143			(+1)	
39 < 212					195 > 96			(-1)		99 < 111			(+1)		183 > 143			(-1)	
60 < 67					138 < 183			(+1)		163 > 96			(-1)		163 < 195			(+1)	
60 > 53					138 < 143			(-1)		163 < 195			(-1)		163 > 96			(+1)	
67 >53					183 > 143			(+1)		96 < 195			(-1)		195 > 96			(+1)	
					(-2)					(-6)					(+6)				

Note. Concordance indices are calculated for two permutations of the ESP matrix to begin computing the overall test statistic for the visual and ESP suit matrices (Kelly et al., 1975).

A closer look at the target set yields more interesting results. Specifically, the 13 × 13 matrix for the visual and ESP number data does show significant concordance. I implemented the within-stimulus triad test to determine the presence of a systematic relationship between the within-stimulus structural properties of confusions in the 13 × 13 visual recognition and ESP number matrices from the Kelly et al. (1975) study. The observed index value is 185 with 458 consistencies and 273 inconsistencies. Out of 6,227,020,800 possible permutations, 27,053,516 had indices greater than or equal to 185. Hence, the one-tailed p value is .004344554, indicating that the visual and ESP number matrices do show fairly impressive similarity in within-stimulus patterning among the off-diagonal (misses) entries.

The analyst should also be familiar with the Mantel (1967) statistic, which provides a framework for determining a one-tailed p value to test the conformity of two matrices defined on the basis of multiplying corresponding matrix entries. The Mantel statistic simply adds the products of corresponding entries, ignoring diagonal entries. A test of agreement (or concordance) between two matrices can be performed by generating a distribution of indices across all possible permutations of the n targets in the second matrix. If computer resources provide the feasibility to evaluate each of the $n!$ permutations of the rows and columns of the matrix, then a complete distribution for the Mantel statistic can be determined and the actual statistic can be mapped to that distribution. Otherwise, a random sampling of permutations can be used to approximate the distribution for the test statistic. A succinct Mantel statistic algorithm can be inserted into the algorithm of the Appendix.

```
Stat = 0
for i = 1:n
        for j = 1:n (i ≠ j)
                Stat = Stat + C1[i, j] * C2[Targets(i), Targets(j)]
end j
end i
return Stat
```

Table 2 shows the Mantel index for the observed data computed for the 4 × 4 visual and ESP suit matrices from Kelly et al. (1975). The index is also computed for the visual matrix and a permutation of the ESP matrix. If the indices are computed for the remaining 22 permutations of the ESP matrix, then we can map the observed statistic to the statistic distribution. Specifically, of the 24 permutations, 18 yield indices that are as large or larger than 132,589. This gives us a one-tailed p value of 18/24 = .75.

TABLE 2
CALCULATING THE MANTEL INDEX TO DETERMINE CONCORDANCE BETWEEN VISUAL AND ESP SUIT DATA

Visual suit (C_{VISUAL})				ESP suit (C_{ESP})				Permuted ESP suit ($C_{ESP(3)}$)			
C	S	H	D	C	S	H	D	D	H	S	C
C	113	29	67	C	175	154	100	D	143	183	138
S 51		23	53	S 213		111	99	H 96		195	163
H 71	39		212	H 163	195		96	S 99	111		213
D 60	67	53		D 138	183	143		C 100	154	175	
				113*175 + 29*154 + 67*100 +51*213 + 23*111 + 53*99 +71*163 + 39*195 +212*96 +60*138 + 67 *183 + 53*143 =117254				113*143 + 29*183 + 67*138 +51*96 + 23*195 + 53*163 +71*99 + 39*111 +212*213 +60*100 + 67 *154 + 53*175 = 130839			

Note. This begins computing of the overall test statistic for the visual and ESP suit matrices (Kelly et al., 1975). The permutation is chosen as (D—H—S—C).

For the 13 × 13 visual and ESP number matrices obtained by Kelly et al. (1975), I used total enumeration for an exact test. The Mantel index for these two matrices is 17381. Of the 6,227,020,800 possible permutations of the 13 × 13 ESP number matrix, only 234,619 permutations yield an index as good as or better than 17381, giving the concordance of the visual and ESP number matrices a one-tailed p value of 234,619/6,227,020,800 = .00003768. Clearly, these two matrices have significant concordance using the Mantel statistic.

Hubert (1978) observed that Mantel-type indices are not necessarily invariant under monotone (order-preserving) transformations of the data. Because the statistics are computed based on one-to-one products of corresponding elements of matrices, different significance level conclusions could be obtained after what might otherwise be considered a standard transformation such as row-normalization or ranking data. This caveat should not be confused with data that are rated (or ranked) by judges. Even with this possible limitation, the Mantel statistic is informative when applied properly. One important application of the Mantel statistic is in uncovering symmetry in confusion matrices.

A Permutation Test for Assessing the Symmetry of a Confusion Matrix

In psychological and parapsychological testing, subjects do not always uniformly cross-confuse targets, due possibly to displacement, lag and residual memory of previous targets. This tendency manifests itself as asymmetry in confusion matrices. Investigating symmetry in an asymmetric matrix might seem counterintuitive. However, because confusion matrices tend to capture confusion as well as statistical noise, uncovering a degree of symmetry is useful for analysts who wish to reveal the most confusion and the least noise. Specifically, a high degree of confusion among pairs of elements can indicate the need to revise choices of targets in future experiments. Concordance, if not equality, between a matrix and its transpose can uncover symmetry in the matrix (Hubert, 1987, p. 196). The analyst is reminded that any matrix can be reduced to the sum of a symmetric matrix and a skew-symmetric matrix (Tobler, 1976).

A permutation test for the symmetry of a confusion matrix can be obtained by calculating the Mantel index for the matrix and its transpose (see Hubert & Baker, 1979 for an extended discussion of this test). To obtain the transpose of a matrix, the rows of the matrix become the columns and vice versa. To illustrate, we applied a symmetry test to the visual recognition and ESP matrices for suits from Kelly et al.'s (1975) study as shown in Table 3. Once again, total enumeration was used to generate the complete distributions for the test statistics. In both cases for suits, the concordance index is relatively low, revealing a lack of symmetry in both matrices with p values that are not significant. This could be indicative of very noisy data in both matrices. Applying this method to the 13×13 matrices for numbers, we see that the one-tailed p value for visual data for numbers is significant, indicating very symmetric tendencies in this matrix. The permutation testing also indicates symmetry in the ESP data for numbers.

A Brief Note on Permutation Tests for Agreement of Three or More Confusion Matrices

To this point, the matrix permutation test has been described within the context of comparing the agreement between two proximity matrices. Hubert (1979a, 1979b) extended this permutation test to three or more proximity matrices. The generation of the complete distribution for the $(n!)^Q$ possible realizations of the index value is impractical for most n and Q, where Q is the number of matrices. This limitation necessitates the reliance on Monte Carlo sampling methods (i.e., using random number generators) to evaluate the significance of the index. If the statistic is sufficiently extreme with respect to the simulated

agreement, then the null hypothesis of no agreement among the matrices is rejected.

TABLE 3
THE MANTEL STATISTIC APPLIED TO MATRICES AND THEIR TRANSPOSES TO DETERMINE PRESENCE OF SYMMETRICAL DATA

	Mantel Index	Number of Permutations Yielding Indices as Good as or Better Than Observed Data	one-tailed *p* value
Visual Suit (4 × 4)	55052	16	.67
ESP Suit (4 × 4)	259334	21	.875
Visual Number (13 × 13)	18294	831474	.00013353
ESP Number (13 × 13)	26320	53189698	.00854176

Note. The original study by Kelly et al. (1975) reported *p* values according to row-normalized data.

OTHER PERMUTATION-BASED METHODS FOR UNCOVERING STRUCTURE IN CONFUSION MATRICES

The preceding permutation tests are designed to examine structure of confusion matrices as a means to understand confusion between targets. Permutation tests analyze confusion (off-diagonal elements) by examining confusion between pairs of targets, regardless of the ordering of rows and columns of the confusion matrix/matrices. Very often, particularly in parapsychology, full examination of confusion needs to engender an understanding of confusion within an entire set of targets. To achieve this goal, we should seek an *optimal* permutation to show a general flow of confusion in the target set (as shown by seriation) and higher degrees of confusion among subsets of the whole target set (as shown by cluster analysis). The succeeding methodologies seek optimal permutations to order the targets in terms of the degree of confusion they engender with other targets. More specifically, we order the targets in such a way that targets are placed "closer" in the sequence when they have a higher degree of confusion relative to the other targets.

The following methodologies are presented with three objectives in mind. First, they are presented to illustrate that permutation-based methodologies are not limited to permutation tests. Second, comments have been made in the parapsychological literature (for example, see May, Utts, Humphrey, Luke, Frivold, & Trask, 1990) that indicate how these methodologies might be useful. Finally, they are presented as mere introductions rather than exhaustive and rigorous explanations.

Seriation

Seriation is designed to obtain a permutation of rows and, simultaneously, columns of one or more confusion matrices so as to more clearly reveal structure among the stimuli. In parapsychology, an optimal permutation will order the targets in such a way as to help explain the relationships that may be present among the targets, or rather, among the confusion/recognition of targets.

This methodology would probably be most useful in parapsychological experiments that atomistically break down descriptions of images or concepts being perceived via anomalous cognition. If such an experiment produced an optimal ordering that could be interpreted as "sensible," then anomalous cognition could be inferred and the experimenter would have a reasonable direction in which to pursue further research. For example, in the transcontinental remote-viewing experiment of Schlitz and Gruber (1980), asymmetric matrices were constructed with data for protocol sites and corresponding transcripts for those sites. Five judges ranked each transcript with each site. (A second matrix was constructed similarly with judges' ratings.) Hubert (1987) characterized the useful applications of seriation in psychology: "For psychologists, the basic interest in the problem of seriation using asymmetric proximities results from the experimental paradigm commonly known as the technique of paired comparisons .... As a slight variation that should be mentioned, each subject could be forced to provide a linear ranking of the n objects." (pp. 137-138). Schiltz and Gruber explicitly stated that their target pool was carefully constructed to contain several targets of given types, which certainly suggests that a seriation according to confusion among those targets could be illuminating.

The *dominance index* is perhaps the most widely used index for seriation of asymmetric matrices (Hubert, Arabie, & Meulman, 2001), with a rich history in the biometric and psychometric literature (Brusco, 2002; Brusco & Stahl, 2001a; DeCani, 1969, 1972; Flueck & Korsh, 1974; Hubert, 1976; Hubert & Golledge, 1981; Rodgers & Thompson, 1992). Essentially, maximization of the dominance index is achieved by finding a permutation that maximizes the sum of confusion elements above the main diagonal. For any pair of targets, the tendency will be to place target i to the left of target j in the sequence if target j is more often mistakenly called for target i than target i is mistakenly called for target j. Lawler (1964) provides a mathematical model and recommends using the optimization process known as dynamic programming to achieve this end.

Alternative seriation goals can determine which criteria an analyst chooses to use to optimally seriate targets/responses. For example, Hubert and Golledge (1981) suggested a nonmetric counting rule, which examines

each column and, for each of the entries above the main diagonal in that column, accumulates the number of times the entry is larger than lower triangle entries in the same column. Notice that this criterion examines within-column (or, for parapsychological confusion matrices, within-response) relationships between entries, somewhat like an upside-down within-stimulus triad test, and seeks to arrange the data in such a way as to show the patterning more clearly.

Aside from choosing the proper criteria for seriation, an analyst must choose an optimization procedure. Because seriation presents combinatorial problems, difficulty increases exponentially as the number of targets increases. Just as the size of the target set often determines whether we use *exact* or *approximate* permutation tests to evaluate degree of matrix structure, the size of the target set determines whether we use *optimal* or *heuristic* methods to uncover matrix structure via seriation. Dynamic programming is the preferred solution procedure when there are roughly 25 or fewer targets in the set (Brusco & Stahl, 2001a; Hubert et al., 2001; Hubert & Golledge, 1981). Branch-and-bound or integer programming methods can sometimes provide optimal solutions for even larger target sets (see Brusco, 2001; DeCani, 1972; Flueck & Korsh, 1974 for extended discussions of these methods). In short, optimal procedures are achieved with total enumeration ($n \leq 13$), dynamic programming ($n \leq 25$), and either branch-and-bound techniques or integer linear programming ($n \leq 35$). However, very large problems require heuristic procedures, which often obtain optimal or near-optimal solutions in a reasonable time frame for problems that would otherwise be intractable.

Unidimensional Scaling Techniques

Seriation can be considered "unidimensional seriation" in that the resulting permutation can be considered placement of stimuli on a number line in an optimal order. Unidimensional scaling goes a step further in that the stimuli are placed along a number line in accordance with their relationship to one another, with greater distance indicating less confusion between the stimuli. Not surprisingly, the seriation methodology in the previous section is fundamental to unidimensional scaling.

A matrix is said to have Robinson structure when values within the matrix decrease moving from the diagonal toward the sides in the rows and moving from the diagonal toward the top/bottom in the columns (Robinson, 1951). For example, looking at the $\mathbf{C}_{ESP}$ matrix in Table 2, the rows C, S, and H have Robinson form, whereas the row D does not. Anti-Robinson structure, the opposite of Robinson structure patterning, is especially important for the purpose of unidimensional scaling. Just as the dominance index is often used to describe asymmetric

matrices, Anti-Robinson structure is generally used to describe symmetric matrices. In fact, for symmetric matrices, perfect Anti-Robinson form is indicative of scalability on a number line.

Cluster Analysis

Cluster-analytic techniques provide another option for recovering structure from confusion data. These methods have not been widely deployed in parapsychological research but there are some noteworthy exceptions. Utts (1993) describes an especially important implementation wherein hierarchical clustering methods were used to form target packets for remote viewing experiments (see also Humphrey, May, & Utts, 1988). Analogous similarity or dissimilarity indices could be constructed from confusion data, and there are a number of techniques that could be applied to partition the targets based on this information. For example, Brusco and Stahl (2001b) recently developed a number of mathematical programming models that can be used to select *subsets* of targets from confusion matrices. Such cluster-analytic partitioning maximizes similarity within clusters and maximizes dissimilarity between clusters, that is, strives to make clusters homogeneous and well separated. These methods are quite flexible and can incorporate information from the main diagonal hits as well as the off-diagonal confusions.

There are important methodological choices, such as the appropriate similarity (or dissimilarity) measure, the appropriate clustering index (just as we would choose an appropriate seriation criterion or permutation test statistic), and the number of clusters and their relative sizes. Therefore, deployment of clustering techniques for confusion matrices is apt to be in experiments designed such that there is a theoretically plausible model for consistent missing, such as a model to examine whether people confuse shapes more often than colors or vice versa. In these situations, a priori hypotheses regarding the number of clusters exist, and thus clustering methods are a natural choice.

I applied a partitioning algorithm to the 13 × 13 visual and ESP number matrices taken from Kelly et al. (1975). The raw confusion data were converted to a symmetric dissimilarity matrix by adding the corresponding matrix elements about the diagonal and then subtracting the sums from maximum across all sums. The objective of the partitioning algorithm was to minimize the maximum dissimilarity between pairs of objects in the same cluster, known as the partition diameter. A six-cluster solution was obtained as shown in Table 4 for the visual and ESP number data from Kelly et al. Half of the clusters are identical and the other half are similar. Although this methodology does not produce the graphic representation as MDS does (as employed by Kelly et al.), the information

is relatively easy to glean and, in this case, revealing in the similarity between the visual and ESP clusters.

TABLE 4
RESULTS OF PARTITIONING CLUSTER ANALYSIS MINIMIZING PARTITION DIAMETER FORTHE VISUAL AND ESP NUMBER DATA FROM KELLY, ET AL. (1975)

Cluster	Visual number data	ESP number data
1	{A, 2, 3}	{A, 2, 3}
2	{4, 5}	{4, 9}
3	{6, 7}	{5, 6, 7}
4	{8}	{8}
5	{9, 10}	{10}
6	{J, Q, K}	{J, Q, K}

DISCUSSION

The study of confusion need not be confusing. By taking advantage of methodologies developed in other fields, parapsychologists can closely examine the often chaotic, noisy, and/or imprecise confusion in psi abilities. The analytical tools presented in this paper—permutation tests, seriation, and cluster analysis—represent some of the well-developed tools for studying confusion.

Most of this methodology is familiar to analytical parapsychologists. Permutation testing and cluster analysis have been used in parapsychological analyses but not fully exploited and not applied to the study of confusion in extrasensory perception. For example, cluster analysis has been used to analyze remote viewing data (May et al., 1990; Utts, 1993), and the counting method on the trace of matrices (Scott, 1972) has also been used to examine remote viewing data (Schlitz & Gruber, 1980). However, the complex, atomistic lists of characterizations of targets in remote viewing trials suggest that seriation might prove more enlightening than sharply cut clusters. In fact, May et al. (1990) suggested "refinement of cluster analysis for targets, in an effort to simulate, as closely as possible, what is meant by 'visual similarity' between targets" and "refinement of the analysis of responses, in an effort to achieve even greater correlations between the fuzzy set figure or merit analysis and various forms of ground truth" (p. 210). This does seem to suggest that seriation, which could provide more of a flowing continuum of similarity/dissimilarity than distinct clusters, would be useful.

Permutation tests are not the ideal methodology for all circumstances. For example, Utts (1989) described some serious problems with Gilmore's (1989) suggestion for using permutation tests for the matching of patterned target sequences. In addition, the results of Rasmussen (1989) and Hayes (1996) are good examples of potential problems with permutation tests when used to

test hypotheses regarding population parameters. Although permutation tests tend to be free of distributional assumptions, some assumptions (e.g., homogeneity of variance) might be required if the objective is to test a parameter. Despite these caveats, permutation tests are broadly applicable and are quite easy to implement. The marked increases in computer processing speed that have arisen since the development of these methods permit complete distribution generation when the number of targets is modest, and also enable a larger number of simulated trials when total enumeration is not computationally feasible. Furthermore, contemporary methodologies exist to enhance permutation-based methodologies in terms of computation feasibility and usefulness in analyses.

I have presented and demonstrated a range of methods using published confusion data from the parapsychological literature. Of course, the selection of criteria for permutation tests should be motivated by the particular research problem at hand. As observed by Kennedy (1979), implementations should be driven by good theory. We echo the sentiment expressed by Burdick and Kelly (1977) more than 25 years ago, and sincerely hope that availability of the procedures described in this article will foster more research regarding confusion structures in ESP data.

References

Alexander, C.H., & Broughton, R. S. (2001). Cerebral hemisphere dominance and ESP performance in the autoganzfeld. *Journal of Parapsychology*, **65**, 408-409.

Blair, R.C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, **30**, 518-524.

Bradley, J.V. (1968). *Distribution-free statistical tests.* Englewood Cliffs, NJ: Prentice-Hall.

Brusco, M.J. (2001). Seriation of asymmetric proximity matrices using integer linear programming. *British Journal of Mathematical and Statistical Psychology*, **54**, 367-375.

Brusco, M.J. (2002). Identifying a reordering of the rows and columns of multiple proximity matrices using multiobjective programming. *Journal of Mathematical Psychology*, **46**, 731-745.

Brusco, M.J., & Stahl, S. (2001a). A multiobjective approach to combinatorial data analysis. *Psychometrika*, **66**, 5-24.

Brusco, M.J., & Stahl, S. (2001b). Compact integer programming models for extracting subsets of stimuli from confusion matrices. *Psychometrika*, **66**, 405-419.

Burdick, D.S., & Kelly, E.F. (1977). Statistical methods in parapsychological research. In B. B. Wolman (Ed.) *Handbook of parapsychology* (pp. 81-130). New York: Van Nostrand Reinhold.

CADORET, R. J. (1957). Note on consistent missing. *Journal of Parapsychology*, **21**, 154-158.

CADORET, R., & PRATT, J.G. (1950). The consistent missing effect in ESP. *Journal of Parapsychology*, **14**, 244-256.

CHO, R.Y., YANG, V., & HALLETT, P.E. (2000). Reliability and dimensionality of judgments of visually textured materials. *Perception & Psychophysics*, **62**, 735-752.

CRANDALL, J.E. (1989). Reinforcement effect and displacement trend: No wine in old bottles. *Journal of Parapsychology*, **53**, 61-67.

DECANI, J.S. (1969). Maximum likelihood paired comparison ranking by linear programming. *Biometrika*, **56**, 537-45.

DECANI, J.S. (1972). A branch-and-bound algorithm for maximum likelihood paired comparison ranking. *Biometrika*, **59**, 131-135.

DOBYNS, Y.H., DUNNE, B.J., JAHN, R.G., & NELSON, R.D. (1992). Response to Hansen, Utts, and Markwick: Statistical and methodological problems of the PEAR remote viewing experiments. *Journal of Parapsychology*, **56**, 115-146.

EDGINGTON, E.S. (1964). Randomization tests. *Journal of Psychology*, **57**, 445-449.

EDGINGTON, E.S. (1966). Statistical inference and nonrandom samples. *Psychological Bulletin*, **66**, 486-487.

EDGINGTON, E.S. (1969). Approximate randomization tests. *Journal of Psychology*, **72**, 143-149.

EDGINGTON, E.S. (1995). *Randomization tests*. New York: Marcel Dekkar.

FISHER, R.A. (1928). A method of scoring coincidences in tests with playing cards. *Proceedings of the Society for Psychical Research*, **34**, 181-185.

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

FISK, G.W., & MITCHELL, A.M.J. (1953). ESP experiments with clock cards: A new technique with different scoring. *Journal of the Society for Psychical Research*, **37**, 1-14.

FISK, G.W., & WEST, D. J. (1957). Towards accurate predictions from ESP data.. *Journal of the Society for Psychical Research*, **38**, 157-162.

FLUECK, J.A., & KORSH, J.F. (1974). A branch search algorithm for maximum likelihood paired comparison ranking. *Biometrika*, **61**, 621-626.

GILMORE, J.B. (1989). Randomness and the search for psi. *Journal of Parapsychology*, **53**, 309-340.

HANSEN, G.P., UTTS, J., & MARKWICK, B. (1992). Critique of PEAR remote-viewing experiments. *Journal of Parapsychology*, **56**, 97-114.

HAYES, A.F. (1996). Permutation test is not distribution-free: Testing H_0: $\rho = 0$. *Psychological Methods*, **1**, 184-198.

HETTINGER, T.P., GENT, J.F., MARKS, L.E., & FRANK, M.E. (1999). A confusion matrix for the study of taste perception. *Perception & Psychophysics*, **61**, 15-21.

HUBERT, L. (1976). Seriation using asymmetric proximity measures. *British Journal of Mathematical and Statistical Psychology*, **29**, 32-52.

HUBERT, L.J. (1978). Generalized proximity function comparisons. *British Journal of Mathematical and Statistical Psychology*, **31**, 179-192.

HUBERT, L.J. (1979a). Matching models in the analysis of cross-classifications. *Psychometrika*, **44**, 21-41.

HUBERT, L. J. (1979b). Generalized concordance. *Psychometrika*, **44**, 135-141.

HUBERT, L.J. (1987) *Assignment methods in combinatorial data analysis.* New York: Marcel Dekker.

HUBERT, L., ARABIE, P., & MEULMAN, J. (2001) *Combinatorial data analysis: Optimization by dynamic programming.* Philadelphia: SIAM.

HUBERT, L.J., & BAKER, F. B. (1979). Evaluating the symmetry of a proximity matrix. *Quality and Quantity*, **13**, 77-84.

HUBERT, L.J., & GOLLEDGE, R.G. (1981). Matrix reorganization and dynamic programming: Applications to paired comparisons and unidimensional seriation. *Psychometrika*, **46**, 429-441.

HUMPHREY, B.S., MAY, E.C., & UTTS, J.M. (1988). Fuzzy set technology in the analysis of remote viewing. *Proceedings of Presented Papers: The Parapsychological Association 31st Annual Convention*, 378-394.

KELLY, E.F. (1980). Further notes on consistent missing. *Journal of Parapsychology*, **44**, 57-61.

KELLY, E.F., KANTHAMANI, H., CHILD, I.L., & YOUNG, F.W. (1975). On the relation between visual and ESP confusion structures in an exceptional subject. *Journal of the American Society for Psychical Research*, **69**, 1-31.

KENNEDY, J.E. (1979). Consistent missing: A type of information-processing error in ESP. *Journal of Parapsychology*, **43**, 113-128.

KENNEDY, J.E., KANTHAMANI, H., & PALMER, J. (1994). Psychic and spiritual experiences, health, well-being, and meaning in life. *Journal of Parapsychology*, **58**, 353-383.

KENT, P.F., YOUNGENTOB, S.L., & SHEEHE, P.R. (1995). Odorant-specific spatial patterns in mucosal activity predict perceptual differences among odorants. *Journal of Neurophysiology*, **74**, 1777-1781.

LAWLER, E.L. (1964). A comment on minimum feedback arc sets. *IEEE Transactions on Circuit Theory*, **11**, 296-297.

LOOMIS, J.M. (1982). Analysis of tactile and visual confusion matrices. *Perception & Psychophysics*, **31**, 41-52.

MAHER, M.C. (1999). Riding the waves in search of the particles: A modern study of ghosts and apparitions. *Journal of Parapsychology*, **63**, 47-80.

MANNING, S.K., & SHOFNER, E. (1991). Similarity ratings and confusability of lipread consonants compared with similarity ratings of auditory and orthographic stimuli. *American Journal of Psychology*, **104**, 587-604.

MANTEL, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209-220.

MAY, E.C., UTTS, J.M., HUMPHREY, B.S., LUKE, W.L.W., FRIVOLD, T.J., & TRAKS, V.V. (1990). Advances in remote-viewing analysis. *Journal of Parapsychology*, **54**, 193-228.

MILTON, J. (1987). A critical review of the displacement effect. *Proceedings of Presented Papers: The Parapsychological Association 30th Annual Convention*, 125-128.

MORGAN, B.J.T., CHAMBERS, S.M., & MORTON, J. (1973). Acoustic confusion of digits in memory and recognition. *Perception & Psychophysics*, **14**, 375-383.

MORRIS, R.L. (1972). An exact method for evaluating preferentially matched free-response material. *Journal of the American Society for Psychical Research*, **66**, 401-407.

PITTMAN, E.J.G. (1937). Significance tests which may be applied to samples from any population: II. The correlation coefficient test. *Journal of the Royal Statistical Society*, **4**, 225-232.

PRATT, J.G., & BIRGE, W.R. (1948). Appraising verbal test material in parapsychology. *Journal of Parapsychology*, **12**, 236-256.

PRATT, J.G., & FOSTER, E.B. (1950). Displacement in ESP card tests in relation to hits and misses. *Journal of Parapsychology*, **14**, 37-52.

RADIN, D.I., MACHADO, F.R., & ZANGARI, W. (1998). Effects of distant healing intention through time and space: Two exploratory studies. *Proceedings of Presented Papers: The Parapsychological Association 41st Annual Convention*, 143-161.

RASMUSSEN, J.L. (1989). Computer-intensive correlation analysis: Bootstrap and approximate randomization techniques. *British Journal of Mathematical and Statistical Psychology*, **42**, 103-111.

RHINE, J.B. (1952). The problem of psi-missing. *Journal of Parapsychology*, **16**, 90-129.

ROBINSON, W.S. (1951). A method of chronologically ordering archeological deposits. *American Antiquity*, **16**, 293-301.

RODGERS, J.L., & THOMPSON, T.D. (1992). Seriation and multidimensional scaling: A data analysis approach to scaling asymmetric proximity matrices. *Applied Psychological Measurement*, **16**, 105-117.

RUSSELL, W. (1943). Examination of ESP records for displacement effects. *Journal of Parapsychology*, **7**, 104-117.

SCHLITZ, M., & GRUBER, E. (1980). Transcontinental remote viewing. *Journal of Parapsychology*, **44**, 305-317.

SCHMIDT, S., SCHNEIDER, R., BINDER, M., BURKLE, D., & WALACH, H. (2001). Investigating methodological issues in EDA-DMILS: Results from a pilot study. *Journal of Parapsychology*, **65**, 59-82.

SCOTT, C. (1972). On the evaluation of verbal material in parapsychology: A discussion of Dr. Pratt's monograph. *Journal of the Society for Psychical Research*, **46**, 79-90.

SOLFVIN, G.F., KELLY, E.F., & BURDICK, D.S. (1978). Some new methods of analysis for preferential-ranking data. *Journal of the American Society for Psychical Research*, **72**, 93-109.

STUART, C.E. (1942). An ESP test with drawings. *Journal of Parapsychology*, **6**, 20-43.

TARG, R. (1994). Remote-viewing replication: Evaluated by concept analysis. *Journal of Parapsychology*, **58**, 271-284.

TIMM, U. (1969). Mixing up of symbols in ESP card experiments (so-called consistent missing) as a possible cause for psi-missing. *Journal of Parapsychology*, **33**, 109-124.

TOBLER, W.R. (1976). Spatial interaction patterns. *Journal of Environmental Studies*, **6**, 271-301.

TOWNSEND, J.T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, **9**, 40 -50.

UTTS, J.M. (1989). Randomness and randomization tests: A reply to Gilmore. *Journal of Parapsychology*, **53**, 345-351.

UTTS, J.M. (1993). Analyzing free response data: A progress report. In L. Coly & D. S. McMahon (Eds.) *Psi research methodology: A re-examination. Proceedings of an international conference* (pp. 71-83). New York: Parapsychology Foundation.

VEGA-BERMUDEZ, F., JOHNSON, K.O., & HSIAO, S.S. (1991). Human tactile pattern recognition: Active versus passive touch, velocity effects, and patterns of confusion. *Journal of Neurophysiology*, **65**, 531-546.

2352 Hampshire Way
Tallahassee, FL 32309
BlueRoseGarden@aol.com

APPENDIX

Systematic Generation of Permutations

Determine the number of targets, *n*.
Set *Index* = 1 and *np* = 1, where *np* is the tally of the total number of permutations. Notice that *np* is initialized at 1 because the first permutation is given as the simple enumeration of targets, 1...*n*.
Set the current position under consideration, *Position* = *n* − 1.

```
for i = 1:n
        Targets(i) = i
        {Enumerate the targets in a vector. Notice that the initial
        enumeration inherently places the target positions from least to
        greatest. These are the initial positions of the targets and the
        assigned values of the targets during the permutation generation
        process.}
        Selection(i) = 1
        {Initialize a vector to be used in the systematic selection of targets 1
        if already selected in a permutation, else 0. In this case, all positions
        are assigned at the start.}
end
Evaluate the Actual statistic.
while Position > 0
        if Targets(Position) > Targets(Position + 1)
                Position = Position - 1
                continue
                {This pushes the targets in the final positions towards the
                beginning of the permutations being generated. If the
                assigned value of the target in Position is greater than the
                assigned value of the target in the next position, then skip
                the "else" and try again.}
        else
                Marker = Targets(Position)
                {The Marker is the target in the position under
        consideration}
                for j = Position:n
                        Selection(Targets(j)) = 0
                        {De-select the targets holding the higher positions.}
                end
                for j = Marker + 1:n
                        if Selection(j) = 0
                                Targets(Position) = j
```

```
                    {Set the target in Position to its next largest
available value.}
                    Selection(j) = 1
                    break
                    {Prevent selection of more than 1 target.}
            end
        end
        Marker2 = Position
        {Marker2 assists in assigning unselected targets to
    unoccupied positions.}
        for j = 1:n
            if Selection(j) = 0
                Marker2 = Marker2 + 1
                Targets(Marker2) = j
                Selection(j) = 1
            end
        end
        np = np + 1
        Evaluate the index given the full permutation of Targets.
        If the generated permutation produced an index more
        extreme than the actual statistic, then increment Index.
        Position = n - 1
        {If Position was decremented at the beginning of the whole
        loop on a previous pass, then reset it to n - 1.}
    end
end
```

Finally, calculate the one-tailed p-value, $p1 = Index/np$.