AD-A269 699

(13)

A REVISED MODIFIED PARALLEL ANALYSIS (RMPA) FOR THE CONSTRUCTION OF UNIDIMENSIONAL ITEM POOLS

David V. Budescu

The University of Haifa
Haifa 31905, ISRAEL

Yoav Cohen **Anat Ben-Simon**

National Institute for Testing and Evaluation (NITE)
POBox 26015, Jerusalem 91260, ISRAEL

July, 1993

NITE RESEARCH REPORT No. 176

This research was sponsored by:

Manpower, Personnel and Training R & D Program
Cognitive Science program Office of Neval Research (ONR)

Under Contract No. N00014-91-J-1666 & T No. 4428034

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704 0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704 0188) Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1993	Final (10/91-4/93)

4. TITLE AND SUBTITLE
A REVISED MODIFIED PARALLEL ANALYSIS (RMPA) FOR THE CONSTRUCTION OF UNIDIMENSIONAL ITEM POOLS

5. FUNDING NUMBERS
N00014-91-J-1666
WU#: 4428034

6. AUTHOR(S)
David V. Budescu
Yoav Cohen
Anat Ben-Simon

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
(1) University of Haifa
Haifa 31905, Israel

(2) National Institute for Testing & Evaluation
POB 26015, Jerusalem 91260, Israel

8. PERFORMING ORGANIZATION REPORT NUMBER
NITE RR-176

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research (ONR)
800 North Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release.
Distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Modified Parallel Analysis (MPA) is a heuristic method for assessing "approximate unidimensionality" of item pools. It compares the second eigenvalue of the observed correlation matrix with the corresponding eigenvalue extracted from a "parallel" matrix generated by a unidimensional and locally independent model.

Revised Modified Parallel Analysis (RMPA) generalizes MPA and alleviates some of its technical limitations. An important and useful feature is a new method for eliminating items which violate the test's unidimensionality. This is achieved by eliminating items, one at a time to determine their contribution to the matrices' eigenvalues.

We propose a test for detecting items with larger impact in the observed data set, and eliminating them. The new method was tested in several simulations in which unidimensional item pools were "contaminated" by various proportions of items from a secondary pool. The results indicate that RMPA does an excellent job in detecting low (10%) and moderate (25%) levels of contamination, but fails in cases of maximal (50%) contamination.

14. SUBJECT TERMS
Parallel Analysis, Dimensionality, Gapping, Unidimensionality, Item Pools

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

A REVISED MODIFIED PARALLEL ANALYSIS (RMPA) FOR THE CONSTRUCTION OF UNIDIMENSIONAL ITEM POOLS

TABLE OF CONTENTS

A REVISED MODIFIED PARALLEL ANALYSIS (RMPA) FOR THE CONSTRUCTION OF UNIDIMENSIONAL ITEM POOLS

BACKGROUND

The increasing popularity of Item Response Theory (IRT) (e.g. Hambleton, 1983; Hulin, Drasgow & Parsons 1983; Lord, 1980) in educational, personnel and psychological testing has caused a revolution in this domain. These new models enable researchers and test users to solve efficiently otherwise intractable problems and develop many innovative testing procedures.

Perhaps the most promising, and undoubtedly the most intriguing, one is Computerized Adaptive Testing (CAT). The basic ideas as well as the theoretical and practical advantages of CAT are well known and widely acknowledged (e.g. Green, 1983; Weiss, 1983). The increasing availability and acceptance of computers in everyday life and their lower prices make CAT a feasible alternative to traditional forms of testing.

Why is it, then, that CAT is relatively slow in replacing conventional testing procedures ? One possible reason are the various problems related to the construction, validation and maintenance of the large item pools required by this new testing protocol.

From a psychometric point of view one of the most interesting and challenging problems is the assessment of the pools' dimensionality. Though multidimensional item response models have been developed (e.g. Reckase, 1985; Sympson, 1978), most readily applicable IRT models used today assume that the test takers responses to all items depend on a single latent trait (ability). Thus, it is crucial to establish that any item used in estimating the examinee's position along this ability continuum measures, in fact, the same trait. In other words, the need to demonstrate that a given item pool is truly unidimensional is a necessary condition for its use in CAT.

Defining and Assessing Unidimensionality

Consider a test consisting of n items selected from a larger item pool. Let $\vec{U}_i$ be the vector of n binary responses to the test's items (taking values of 1 and 0 for correct and incorrect response, respectively), generated by the *ith* test taker (i=1...N), and let U_{ij} be her response to the *jth* item (j=1...n). Finally, let $\vec{\theta}_i$ be a vector of t latent traits characterizing the examinee's abilities. The strong principle of local independence (McDonald, 1981) states that:

$$P(\vec{U}_i = \vec{u}_i | \vec{\theta}_i) = \prod_{j=1}^{n} P_j(U_{ij} = u_{ij} | \vec{\theta}_i) \qquad (1)$$

This principle asserts that the responses to any pair of items are statistically mutually independent for any individual, or any subpopulation with fixed latent traits. The dimensionality of $\vec{U}$ is, simply, the minimal number of latent traits necessary to produce a (strong) locally independent model for $\vec{U}$. Thus, a pool is unidimensional if responses to all its items can be produced by unidimensional locally independent models.

Although a voluminous literature exists on the issue of unidimensionality of items and tests (see Berger and Knol, 1990; Hattie, 1984 and 1985 for partial reviews), currently there is no single approach which is fully satisfactory and/or universally accepted. Hattie (1984) compiled a list of 87 measures of unidimensionality and classified them into five nonoverlapping classes according to their underlying rationale. He distinguished between indices based on

(i) closeness to specific answer patterns,
(ii) reliability coefficients,
(iii) principal components (PC),
(iv) factor analysis (FA) and
(v) goodness of fit to various IRT models.

Hattie questioned the theoretical rationale of indices based on response patterns and reliability and showed empirically that the measures based on PC, FA and one parameter IRT (the Rasch model) are outperformed by methods quantifying deviation from multi-parameter IRT models.

"Approximate" Unidimensionality

Many researchers have argued, based on theoretical and empirical observations, that purely unidimensional tests, or pools, are quite rare (e.g. Ackerman, 1989; Humphreys, 1985; Reckase, Ackerman & Carlson, 1988; Traub, 1983; Yen, 1984, 1985). If, in fact, unidimensionality is frequently violated it is important to determine the practical implications of such violations. Following Reckase's original work (1979), several researchers (e.g. Drasgow & Parson, 1983; Yen, 1984, 1985) have shown that unidimensional models are quite robust under multidimensionality as long as

(i) There is a single "dominant" factor, and
(ii) Item difficulty is not confounded with dimensionality.

These, and other similar, studies suggest that strict unidimensional pools are not necessary for many practical applications of unidimensional IRT models (e.g. CAT). It is, however, important to develop methods that can identify pools which are almost / practically / approximately unidimensional (i.e. they deviate from strict unidimensionality to a degree which does not seriously affect the fit or accuracy of the unidimensional IRT model).

This is the motivation behind recent work by Stout, who developed a test of the *essential unidimensionality* of a data set (Stout, 1987,1990; Nandakumar, 1991). Essential independence is achieved if the mean covariance (conditional on $\vec{\theta}_i$, the test taker's vector of t latent traits) between all n(n-1)/2 pairs of items approaches 0 as the number of items increases to infinity, and the essential dimensionality of a pool is the smallest number of latent traits necessary to satisfy essential independence. Essential independence is a weaker requirement than strong local independence and, in practice, it is obtained whenever there is a single dominant dimension in the data (e.g. Nandakumar, 1991).

In the same spirit Drasgow and Lissak(1983) presented Modified Parallel Analysis (MPA for short) as "a technique that can determine when an item pool is *sufficiently unidimensional* for the use of IRT" (Drasgow and Lissak, 1983, page 365). Modified Parallel Analysis relies on FA, a well understood method which is widely available to users in most statistical packages. Thus, it is (conceptually and computationally) easier to use than Stout's methods. This study will develop a revised and improved version of MPA.

Parallel and Modified Parallel Analysis

Parallel Analysis (PA) was proposed by Horn (1965) as an alternative to traditional factor analytical methods for identifying the number of latent factors. The standard methods are based on various functions of the eigenvalues of the correlation matrix. Among them, the eigenvalues' absolute size (e.g. Kaiser, 1960), their overall pattern (e.g. Cattell, 1966), or their distribution under the multivariate normal model (e.g. Bartlett, 1950).

The rationale behind PA is intuitively compelling, and its application is simple and straightforward: Random correlation matrices are generated, and their eigenvalues are extracted and averaged. The eigenvalues of the actual correlations are compared to these means and those factors with eigenvalues larger than their counterparts from the randomly generated data are retained. Crawford and Koopman (1973), Humphreys and Montanelli (1975) and Zwick and Velicer (1986), among others, report that PA works well in both Principal Components (PC) and Factor Analysis (FA). Recently Longman, Cota, Holden and Fekken (1989) published regression equations that eliminate the need to actually generate random matrices for each PA (for the PC case).

Parallel Analysis is used to determine the true dimensionality of a given data set, whereas in most applications of CAT one seeks to determine whether a data set deviates significantly from unidimensionality. Modified Parallel Analysis (Drasgow & Lissak, 1983) provides an ingenious way of answering this question, using the rationale of PA. Its basic stages are:

(1) The intercorrelations (preferably tetrachoric) of the test's items are factor analyzed and the eigenvalues of the unrotated solution are calculated.

(2) A "parallel" unidimensional data set is generated by an IRT model. This data set parallels the observed one along all its attributes: It has an equal number of examinees with identical abilities, and it has the same number of items with identical parameters. Since responses are generated by an unidimensional IRT model satisfying the strong local independence principle the data set is, by definition, unidimensional.

(3) The (tetrachoric) correlations of the parallel data set are factor analyzed, and the eigenvalues of the unrotated solution are calculated.

(4) The dimensionality of the pool is assessed by comparing the magnitude of the second eigenvalues of the two data sets: If the actual value (calculated in stage 1) is "sufficiently close" to the one obtained from the parallel data set (calculated in stage 3), the test is unidimensional.

Drasgow and Lissak (1983) recommend that the items' communalities be estimated by the largest (absolute) off-diagonal correlation, and suggest an ad hoc procedure for imputation of tetrachoric correlations for those cases where the regular algorithm fails to converge. They also report five empirical studies providing strong empirical support for the procedure.

Eigenvalue based factor analytical techniques are not always successful in recovering the true dimensionality of binary data and, consequently, can't always distinguish between unidimensional and multidimensional data sets (e.g. Collins, Cliff, McCormick and Zatlin, 1986; Hattie, 1984; Knol and Berger, 1991; Roznowsky, Tucker & Humphreys, 1991; Zwick and Velicer, 1986). Thus it may seem surprising that some of the same measures perform very well in the framework of PA, and MPA. It is important to stress that the key to the success of these methods is their comparative nature. Whatever deficiencies these statistics have, they affect equally the results of the two data sets. Both PA and MPA focus on, and highlight, whatever differences exist between the empirical and parallel data sets above and beyond the systematic biases that the FA based measures may share.

Thus, in Hattie's (1984) typology MPA should not be considered a "factor analytic approach". In fact, it is closer to the "measures of fit to IRT models". MPA is a general method for assessing the similarity, or closeness, between two parallel data sets (one of which is known to be unidimensional), in which the similarity is quantified by some of the statistics usually employed in FA.

A critique of MPA

Modified Parallel Analysis suffers from a few technical limitations. In this section we describe these limitations and the problems they may cause in applying the method:

(i) MPA is a randomized procedure, i.e. its results depend to a certain degree on a random process which is totally unrelated to the process of interest, namely, the selection of the parallel data set. Thus, with small enough samples, researchers applying exactly the same procedure to the same set of data may reach different conclusions because of the variance between the random data sets generated in their simulations.

(ii) The simulated and the empirical data sets are equated along most important dimensions and any discrepancy between their eigenvalues can ,supposedly, be attributed to the multidimensionality of the empirical matrix. Yet, the communalities are estimated in a purely empirical fashion separately for each data set, introducing another important difference between them. This factor may bias (in an unknown direction and to an unknown degree) the comparative analysis.

(iii) MPA is a heuristic procedure, i.e. it lacks a measure of sampling variability for the formal assessment of the closeness of the critical statistic (the second eigenvalue) obtained from the unidimensional and the empirical solutions.

Other important limitations of MPA are:

(iv) It compares only the second pair of eigenvalues of the two matrices. This choice lacks a solid theoretical or empirical justification, and it may miss differences between the other eigenvalues (especially the third).

(v) MPA is too limited in its scope. The technique provides a global omnibus test of the hypothesis concerning the pool's unidimensionality. It lacks, however, a mechanism to follow up rejections of the hypothesized pattern, by eliminating some items and identify a unidimensional subset of the pool.

A REVISED MODIFIED PARALLEL ANALYSIS (RMPA)

In this section we outline a revised procedure (RMPA) which extends and generalizes the MPA. The revised method offers solutions to the technical problems described above and incorporates them into the existing framework of MPA. Originally, MPA was developed as a global procedure that distinguishes between (essentially) unidimensional tests and multidimensional ones. RMPA complements this aspect by a second stage which allows one to extract unidimensional subsets from larger, potentially multidimensional, pools.

To solve the first problem we replace the random generation of a parallel unidimensional population by the theoretical derivation of the expected correlations under the assumptions of (1) local independence, (2) unidimensionality of the parameter space and (3) the three parameter logistic model (e.g. Lord, 1980). The probability of a correct response for item j by a test taker with (a single) ability θ_i is given by $P(U_{ij} = 1|\theta_i)$ or, in a shorter notation, P_{ji}:

$$P_{ji} = c_j + \frac{1 - c_j}{1 + \exp\{-1.7a_j(\theta_i - b_j)\}} \tag{2}$$

where a_j is the item's discrimination parameter, b_j is the item's difficulty and c_j is its pseudo-guessing probability (see Hambleton, 1983 or Lord, 1980 for details).

Under these assumptions the expected number of correct answers to any pair of arbitrary items, j and k, in a random sample of N examinees is:

$$f_j = \sum_{i=1}^{N} P_{ji} \text{ and } f_k = \sum_{i=1}^{N} P_{ki} . \tag{3}$$

Under the assumption of local independence, the expected number of correct answers to *both* items, j and k, is:

$$f_{jk} = \sum_{i=1}^{N} P_{ji}P_{ki} . \tag{4}$$

Given f_{jk} and the two marginals, f_j and f_k, the expected 2x2 contingency table can be constructed, and the expected tetrachoric correlation can be estimated by standard methods (e.g. by solving a polynomial using the Newton Raphson method, as suggested by Kendall & Stuart 1979, pages 324-327). All expectations are (as in the original MPA) conditional upon the abilities and item parameters shared by the two data sets. The calculation can be further refined when the true distribution of the unidimensional abilities (θ_i) in the population is known. In these cases, the summation is replaced by integration across all values of θ_i weighted according to the probability density of the θ_i, to yield the matrix of expected tetrechoric correlations in the population.

To solve the second problem we replace the separate estimation of the communalities in the two data sets by the expected tetrachoric correlation between (hypothetical) experimentally independent administrations of any item under the assumptions of (1) local independence, (2) unidimensional ability and (3) a three parameter logistic item curve. This procedure amounts to estimating the items' communalities by their expected test-retest reliabilities. It is well known (e.g. Lord & Novick, 1968; Mulaik, 1972) that a measure's reliability provides an upper bound to its communality. The estimation procedure is just a special case of the technique described above for the calculation of the expected correlation. More specifically, if we let j=k, Equation 4 is reduced to:

$$f_{jj}= \sum_{i=1}^{N} P_{ji}^2 . \qquad (4a)$$

The solution of the third problem relies on a data analytic procedure known as "jacknifing" (see Arvesen and Salsburg, 1975, Miller, 1974 or Mosteller & Tukey, 1977 for partial reviews) [1]. Assume that the original nxn correlation matrix between the test's items is strictly unidimensional. By eliminating one item at a time (i.e. deleting a row, and the corresponding column, from the original matrix) we obtain n submatrices of order (n-1)x(n-1) which, by definition, are also unidimensional. Furthermore, it is easy to show that under the "one factor model" (i.e. a matrix of rank one), the average first eigenvalue of these n submatrices, scaled by a factor of n/(n-1), is an unbiased estimate of the first eigenvalue of the original intact matrix.

An useful and important consequence of the "eliminate one item at a time" procedure is that it provides a simple method for assessing the impact, or influence[2], of any single item on the test's eigenvalues. The logic of the MPA procedure predicts that, under unidimensionality, the two matrices will have equal eigenvalues. For example, it is generally accepted, and it was confirmed empirically by Drasgow & Lissak(1983), that the first eigenvalue (λ_1) is approximately equal in the observed and the expected matrices, regardless of the dimensionality of the observed responses. Thus, except for sampling error, the ratio of the two eigenvalues, RL_1, should be:

$$RL_1 = \lambda_1(\text{observed})/\lambda_1(\text{expected}) = 1 \quad . \tag{5}$$

Furthermore, under unidimensionality, the eigenvalues of the n submatrices of the two data sets will be similar, will have equal variances and will be highly correlated. Finally, the removal of any given item from the pool will affect the observed and the expected data sets in identical fashion and to an equal degree. Thus, equality (5) should also hold in all n submatrices obtained by eliminating one item at a time. Let λ_1^{-i} be the first eigenvalue of the submatrix obtained after the deletion of item i, and let RL_1^{-i} be the ratio of the eigenvalues from the two parallel data sets. Then, for all items (i=1...n), the ratio of the jacknifed eigenvalues should equal the ratio of the original values:

$$RL_1^{-i} = \lambda_1^{-i}(\text{observed})/\lambda_1^{-i}(\text{expected}) = RL_1 \quad . \tag{6}$$

If the responses are unidimensional, similar results are expected to hold for the second, third, and all subsequent eigenvalues. If, on the other hand, the observed responses violate unidimensionality, the analysis of the two data sets should yield differential results. For example, Drasgow and Lissak(1983) based the original MPA on the prediction that the second eigenvalue of the observed matrix will be larger than its counterpart from the parallel unidimensional data set:

$$RL_2 = \lambda_2(\text{observed})/\lambda_2(\text{expected}) > 1 \quad . \tag{7}$$

If the data are generated by a multidimensional model we expect the mean of the n second eigenvalues extracted from the observed submatrices to be larger, and their variance to be higher, than their counterparts from the expected data set. Depending on the type and degree of deviation from unidimensionality, the correlation between the observed and expected values can be low (or even negative). Furthermore, the eigenvalues of the observed responses will be more sensitive to the removal of the foreign (or "contaminating") items. Since the expected matrix is unidimensional, its eigenvalues should not be affected considerably when any arbitrary item is removed. However, when a contaminating item is removed from a multidimensional test, the data set becomes closer to unidimensionality and its eigenvalues should decrease. For example, in a test of length n=50 with 8 foreign items (8/50=16% contamination), after the removal of such an item, the level of contamination is reduced to (7/49=)14%. Thus, whenever a contaminating item is eliminated the matching eigenvalues should be more similar to each other than in those cases in which a regular (noncontaminating) item is removed. Consequently, the ratio of the eigenvalues should be closer to unity in these instances.

To summarize, for any given data set, the ratio between the first eigenvalues, RL_1, in the two data sets can be used as a benchmark against which one can assess and test the ratios derived from the second and third eigenvalues (RL_2 and RL_3, respectively). At the global (i.e. test or pool) level, this approach is attractive because the behavior of RL_2 and RL_3 is assessed by a data based index which is more sensitive to, and reflects, the peculiarities and idiosyncrasies of the specific test being examined. At the local (i.e. item) level, this procedure provides a natural way of ranking, and scaling, the items according to their deviation from the pattern expected under unidimensionality. These properties can be used to develop a procedure for testing the global dimensionality of the observed responses, and a method of selecting unidimensional pools. In the next section we describe the technical details of such a testing procedure.

The "gap test"

As described above, we propose to jacknife the two parallel correlation matrices and calculate the eigenvalues of all n submatrices. To facilitate the comparison of the two data sets we calculate, for all items (i=1...n) and for the first k eigenvalues (typically k=1,2,3 should suffice), the ratio of the two matched eigenvalues:

$$RL_k^{-i} = \lambda_k^{-i}(\text{observed}) / \lambda_k^{-i}(\text{expected}) \quad . \qquad (8)$$

The global ratio RL_1, as well as the individual RL_1^{-i} (i=1..n), are insensitive to the dimensionality of the observed data set. Their empirical distribution will be used to test the hypothesis that the ratios of the second and third eigenvalues behave similarly. Formally, we wish to test that $F\{RL_2^{-i}\} = F\{RL_1^{-i}\}$, and $F\{RL_3^{-i}\} = F\{RL_1^{-i}\}$, where $F\{\cdot\}$ stands for the distribution of the relevant statistic. The alternative hypothesis is that the ratios are distributed differentially.

We are particularly interested in the case where an essentially unidimensional data set is contaminated by a second (sometimes called "nuisance") ability. We speculated earlier, that removal of such contaminating items will affect differentially the two matched eigenvalues. When analyzing the correlations from the observed responses we expect to observe two distinct clusters of eigenvalues --- from the unidimensional and the contaminating pool, respectively --- separated by a substantial "jump". No parallel clustering and separation is expected in the corresponding eigenvalues of the matrix of expected correlations.

To detect such unusual jumps we adopt a procedure described by Wainer and Schacht (1978) under the name of "gapping" since its goal is to detect unusually large gaps in strings of ordered values. The first step in this procedure is to rank order the values in descending order and to calculate the (n-1) gaps, g_i, by subtracting each observation from the immediately previous (i.e. larger) one. The gaps are then weighted by a set of logistic weights to yield weighted gaps, y_i. These weights were selected to account and compensate for the fact that, typically, observations are more dense (hence should be overweighted) near the center and more sparse (and should be underweighted) in the tails of the distribution. Formally:

$$y_i = \sqrt{i\,(n - i)}\, g_i \quad . \tag{9}$$

Finally, these values are standardized by division by y_m, the midmean (i.e. the mean of the central 50% values) of the weighted gaps. Thus, the standardized weighted gaps (SWGs for short), z_i can be expressed as:

$$z_i = y_i / y_m . \tag{10}$$

Zero gaps indicate that two adjacent observations are equal, and unit gaps indicate that the distance between two observations is equal to the gaps' midmean. By definition, all gaps are non-negative but are unbounded from above. Wainer and Schacht (1978) suggest that z_i values greater than 2.25 indicate "unusually" large gaps. The probability of observing gaps this wide by chance is approximately 0.03 under the normal distribution, but this value was shown by Wainer and Schacht (1978) to work quite well for a variety of symmetric t distributions with tails larger than the normal.

We will use this procedure to detect the location of the gap separating the items from the two pools, on the basis of ratios of the matched eigenvalues, $RL_k^{\cdot i}$ ($k > 1$). Thus, the hypothesis will be tested by comparing $MAX(Z_{ki})$, the largest SWG, with a critical rejection threshold. However, in the absence of precise information regarding the form of the distribution of these ratios, and the multiplicity of tests involved, it is not sufficient to rely on the 2.25 universal rule of thumb proposed by Wainer and Schacht. Instead, we find it necessary to develop more general (and more conservative[3]) rejection rules .

There are various ways of deriving critical rejection points for this decision: If the distribution of RL_1^{-i} is known (e.g. normal), the critical values can be obtained from the appropriate table. Otherwise, one can estimate the desired percentiles (.01, .05, etc.) from the distribution of RL_1^{-i}. Finally, one can use a version of Chebyshev inequality (e.g. Stuart and Ord, 1987, page 110). The regular Chebyshev inequality states that the probability of finding a value located more than K standard deviations (SDs) from the population's mean is smaller than $1/K^2$, for any distribution with finite moments; A tighter version, invoking the additional assumptions that the distribution is symmetric and unimodal, yields a lower upper limit ($4/9K^2$), for the probability of the same event [4].

The decision, to reject H_0, will be based on a comparison with a critical threshold, $T(z_1)$. The threshold is derived from the distribution of the ratios of the first eigenvalue, RL_1^{-i}, in the same data set. Specifically, for k=2,3 we will reject H_0 if:

$$MAX(Z_{ki}) > T(z_1) = (M_1 + KS_1)$$

where M_1 and S_1 are the mean and SD, respectively, of the SWGs, z_{1i}, calculated from the ratios of first set of matched eigenvalues, RL_1^{-i}. For the three possible distributional assumptions described above, and with probability of Type I errors fixed at 0.01, 0.05 and 0.10, K takes the values described in the following table:

	Prob (Type I error)		
Assumption	0.01	0.05	0.10
Normality	2.50	2.00	1.65
Symmetry + unimodality	6.67	3.00	2.11
None	10.00	4.50	3.15

The normal case is fully consistent with Wainer and Schacht's 2.25 universal rule of thumb, and needs no further elaboration. It is included in the table, primarily, as a benchmark against which the more conservative Chebyshev rules can be evaluated. We will have more to say about the various rejection rules later in the paper.

A ten - step summary of RMPA

(1) Following the administration of a test consisting of n items to a sample of N test takers, estimate

(i) the three parameters of each item,

(ii) the ability of each examinee, and

(iii) the nxn matrix of tetrachoric correlations between the test's items.

(2) Using the ability and item parameters estimated from the observed responses, calculate the nxn matrix of expected tetrachoric correlations between the items.

(3) The (unit) diagonal values of the observed and expected correlation matrices are replaced by the expected item test-retest reliabilities, and the first k (k=1,2,3) eigenvalues of the two matrices are extracted.

Except for a few technical refinements the previous steps are identical to, and allow the application of, MPA.

(4) Jacknife both correlation matrices by removing one item (row and corresponding column) at a time, and extract the first k (k=1,2,3) eigenvalues of all the (n-1)x(n-1) submatrices.

(5) The corresponding eigenvalues of the observed and expected submatrices are matched and k ratios (k=1,2,3) of the form:

$$RL_k^{-i} = \lambda_k^{-i}(\text{observed}) / \lambda_k^{-i}(\text{expected}) \qquad (8)$$

are calculated for each item (i=1...n).

(6) The n ratios in each of the k sets are rank ordered, SWGs are calculated (Wainer and Schacht, 1978), and the largest SWGs, $MAX(z_{ki})$, are identified.

(7) Using information (Mean, SD, test of normality, etc.) from the distribution of the SWGs based on the first set of matched eigenvalues determine $T(z_1)$, the critical threshold for detecting unusually wide gaps (supposedly distinguishing between items from the primary and contaminating pools).

(8) Compare $MAX(z_{2i})$ and $MAX(z_{3i})$, the largest SWGs based on the second and third set of matched eigenvalues, with $T(z_1)$ the critical rejection threshold.

(9) If $MAX(z_{2i})$ and/or $MAX(z_{3i}) > T(z_1)$, i.e. there is a significant gap in either distribution of ratios, eliminate those items which are located above the significant gap(s)[5].

(10) Let m_1 denote the number of items eliminated ($m_1 > 0$) after this first pass through the data. Repeat stages 4 - 9 with the reduced $(n-m_1)\times(n-m_1)$ correlation matrices. This second analysis may lead to the elimination of additional (say m_2) items. Repeat the procedure with the remaining items, and stop when the test (step 8) fails to detect items to be rejected.

AN EMPIRICAL STUDY OF RMPA

Method

In this section we report results of an empirical study designed to test RMPA. Like most other studies in this area we simulated artificial test results by combining real item parameters and a set of reasonable assumptions regarding the distribution of abilities in the population of test takers. For the purpose of this study we contaminated a large unidimensional pool by (various proportions of) responses generated by a second (nuisance) ability correlated (at various levels) with the first. The efficiency of the RMPA was assessed by its ability to identify correctly the contaminating items and, consequently, partition the test into its two basic components.

We expect this procedure to be most efficient in cases of approximate unidimensionality. In other words, it should detect accurately relatively low levels of contamination, but not mixtures of two (equal) abilities. We also predict that the accuracy of the detection will be inversely related to the correlation between the two abilities involved.

Design

We generated 20 distinct "artificial tests". The following characteristics were fixed for all the tests:

n = test length = 80 items;
N = sample size = 2000 examinees;
t = number of abilities = 2.

The following variables were manipulated across tests:

p = proportion of contaminating items = 0%, 10%, 25% or 50% (p=0% is a a strictly, uncontaminated, unidimensional test and the other three cases represent low, medium and high levels of contamination);

r = the correlation between θ_1 and θ_2, the two abilities = 0.0, 0.5, 0.7 (the three values are approximately equally spaced in terms of r^2).

Replications: All combinations of p and r were replicated twice (i.e. with different seeds for the generation of the abilities, and different item parameters). In the sequel the two replications are labeled "B" and "R".

This design is summarized in the 10 cells of the following table. With the exception of the control condition (p=0, r=0), this can be viewed as a factorial crossing of two independent variables repeated twice.

r=correlation between abilities	p=% of contamination 0	10	25	50
0.0	X	X	X	X
0.5	-	X	X	X
0.7	-	X	X	X

Item Parameters

The items for half the tests (replication "R") were randomly selected from the item bank of a test of English as a Foreign Language (EFL). This test was developed and is routinely used by the National Institute for Testing and Evaluation (NITE) as part of the Psychometric Entrance Test (PET) which is administered to all applicants to universities in Israel. The item parameters were estimated under the three parameter logistic model (Equation 2) using responses from approximately 7,000 examinees who took the test in 1988. The estimation was performed using the NITEST parameter estimation software (Cohen & Bodner, 1989). These parameter estimates for the n=80 items will henceforth be referred to as "true parameters". They are listed in Appendix 1.

The items for the other tests (replication "B") were generated artificially, according to some distributional assumptions: The discrimination parameters (a's) were sampled from a normal distribution with a mean of 1.1 and a s.d. of 0.3; The difficulty parameters (b's) were obtained from a normal distribution with a mean of 0 and a s.d. of 0.8; The pseudo-guessing parameters (c's) are taken from a uniform distribution over the range 0.1 - 0.3. The values of the three parameters were sampled, from the respective sources, independently. The "true parameters" of the "B" tests are listed in Appendix 2.

Table 1 summarizes the information regarding the two sets of true parameters. The two tests are equally difficult, but vary with respect to other aspects. The discrimination parameters of the real items ("R") have a higher mean and variance (m_a=1.33 and s_a=0.51) than the artificial ones ("B") (m_a=1.12 and s_a=0.25). On the average, it is easier to guess in the artificial test (m_c=0.2 vs. 0.16). Finally, whereas the parameters of the artificial items are uncorrelated (by design), the values of the EFL items parameters are moderately correlated.

Insert Table 1 about here

Abilities

All samples include N=2000 simulated "respondents". First we generated four mutually uncorrelated sets of abilities (T, A_1, A_2 and A_3): We sampled 8000 independent observations from the standard (0,1) normal distribution and randomly assigned them to the four sets. Correlated abilities were generated by calculating:

$$T(r) = r \cdot T + \sqrt{1 - r^2} \cdot A_i \qquad (11)$$

where A_i stand for A_1, A_2 or A_3, and r is the desired correlation (0.0, 0.5, 0.7) between the new set of abilities, T(r), and the reference set, T. Thus T(0), T(.5), T(.7) are sets of N=2000 normally distributed abilities which correlate 0.0. 0.5 and 0.7, respectively, with T.

Responses

Four sets of unidimensional response vectors were generated. Each set was simulated with a different set of abilities {T, T(0), T(.5) or T(.7)}, and all responses were generated with the "true" item parameters. The response vectors were simulated with the NITECAT software package (Cohen, Bodner & Ronen, 1989), which implements the process described by Drasgow and Lissak (1983).

The vectors generated with the T abilities are considered the "original" responses based on the dominant ability. Contaminated responses were obtained by replacing the original responses on p% of the items (randomly selected) with the corresponding responses generated by one of the other samples of abilities. Note that for the case of r=0 this procedure simulates a two-dimensional "noncompensatory" model (e.g. Ackerman, 1989, Sympson, 1978), whereas the other cases (r > 0) simulate "compensatory" models (e.g. Ackerman, 1989, Reckase, 1985).

Parameter estimation

In each of the artificial tests the three parameters of the n=80 items were estimated with the NITEST program (Cohen & Bodner, 1989). These are the various sets of "estimated parameters", to be used in the generation of the expected correlations.

Consistent with the massive literature on this topic (e.g. Dorans & Kingston, 1985; Miller & Oshima, 1992; *Oshima & Miller, 1992), we found that the estimates of the* b's *and* c's *were* not affected by the contamination. However, the estimates of the a's (the discrimination parameters) are sensitive to the level of contamination. Appendix 3 presents the mean estimates of the a parameters for items loaded on the dominant and nuisance trait. The pattern and magnitude of the estimates is consistent with other studies in the literature: The estimates for items loaded on the dominant ability are hardly affected, whereas the discrimination measures of the contaminating items are reduced considerably. The magnitude of this "shrinkage" is related to the level of contamination and the correlation between the two factors. A very similar pattern is observed when comparing communality estimates (expected test-retest reliabilities of the items). The results of this comparison are summarized in Appendix 4.

Results

The data were analyzed according to the ten steps procedure outlined in the summary above. We report the main results according to the same sequence.

Standard MPA

At the conclusion of the third stage one can perform the standard MPA, prescribed by Drasgow and Lissak (1983). Table 2 summarizes these results. The table displays the first three eigenvalues of both correlation matrices, as well as their ratios.

Insert Table 2 about here

There is a clear and consistent pattern in the data which can be summarized by three observations:

(i) The first eigenvalues are, practically, equal in the two matrices and their ratio is, essentially, 1. There are no discernible differences between the 18 contaminated data sets and, in this respect, they are indistinguishable from the two uncontaminated tests.

(ii) In all contaminated tests, the second eigenvalue of the observed matrix is larger than its expected counterpart. Consequently, their ratio is greater than unity, as predicted by Drasgow & Lissak (1983). The ratio is a monotonically increasing function of p, the level of contamination, and a monotonically decreasing function of r, the inter-ability correlation.

(iii) The ratio of the third pair of eigenvalues is also greater than one. In fact, in most cases it is greater than the second ratio. The third ratio is not systematically related to r, the inter-ability correlation. However, it increases monotonically as a function of p, the level of contamination. The sharpest effect is obtained for highly ($r=0.7$) correlated, and the weakest effect is found for uncorrelated ($r=0.0$) abilities.

RMPA

At the conclusion of the fifth stage one can perform an informal RMPA by examining the eigenvalues of the jacknifed parallel matrices. Table 3 displays means, and standard deviations, of the first three eigenvalues extracted from the jacknifed submatrices. All the values in the table are based on $n=80$ matrices of order $(n-1)=79$. Note that the mean values are related to the eigenvalues from table 2 through multiplication by a scale factor of $n/(n-1)=$ 80/79.

Insert Table 3 about here

Table 4 presents ratios of the means, and the variances, of the three jacknifed eigenvalues of the 20 tests.

Insert Table 4 about here

There is a close correspondence between these mean ratios and the ratios presented in table 2, and the same three basic conclusions apply here, as well. The ratios of the variances follow a similar, but not identical, pattern:

(i) The variances of the first eigenvalues are, on the average, very close to each other and their ratio is close to unity. The only exceptions are the cases {r=0, p=50}, which represent mixtures of two unidimensional half-tests involving uncorrelated abilities.

(ii) In most cases (and on the average) the variance of the second (jacknifed) eigenvalues in the observed matrices is higher than in the expected one. The effect is most pronounced in the case of the independent traits (r=0), and for moderate or high levels of contamination (p=25 and 50, respectively).

(iii) In all 20 tests the variances of the third (jacknifed) eigenvalues are substantially higher in the observed matrices. The effect is much stronger than for the second eigenvalue, but there is no systematic pattern of change across levels and types of contamination.

Table 5 presents the correlations between the matched jacknifed eigenvalues for the 20 tests. Each correlation is based on n=80 observations.

Insert Table 5 about here

The pattern of results is clear and consistent with our expectations:

(i) There is a high (almost perfect) linear correlation for the first eigenvalue in most tests. The single exception is the {Rep=R, r=0, p=50} case, which is a mixture of two uncorrelated (unidimensional) half-tests.

(ii) In all cases of moderate and high contamination (p=25 and 50, respectively) the correlations based on the second and third eigenvalue are low, or negative.

(iii) In most cases of low contamination (p=10) the correlations based on the second eigenvalue are high (almost like for the first eigenvalue), but the correlations based on the third eigenvalue are always low, or negative.

This pattern indicates that, as suggested by Drasgow and Lissak (1983) and others, the first eigenvalues of the two parallel matrices are practically indistinguishable, across all types and levels of contamination. However, contrary to Drasgow and Lissak's speculation, not all the differences between the two data sets can be detected by comparing the second pair of eigenvalues. The means, variances and correlations of the jacknifed values seem to suggest that in some cases of low contamination ($p=0.10$) violations from unidimensionality can only be detected by examining the third pair of eigenvalues.

Rejection Thresholds

Table 6 presents seven rejection thresholds calculated from the distribution of the first ratio in the 20 tests. The first is, simply, the 2.25 value proposed by Wainer and Schacht (1978). The other six are obtained by crossing two confidence levels (95% and 99%) with three rules of detection --- an empirical value, a value calculated by the "tight" (i.e. assuming unidimodality and symmetry) Chebyshev inequality, and a value derived from the unconstrained Chebyshev inequality.

Insert Table 6 about here

In all tests, and for both confidence levels, the empirical percentile is more liberal than the corresponding Chebyshev bounds. Thus, the three rules can be ranked, from the most to the least conservative, identically for all tests and for both levels of confidence:

Unconstrained Chebyshev > Constrained Chebyshev > Empirical

The 2.25 value is, in all cases, more extreme than the empirical 95th percentile, but smaller than all the Chebyshev bounds. In most cases (13/20) the 99th empirical percentile is above 2.25. One remarkable and reassuring aspect of this table is the relatively low variance of the bounds across the various conditions and replications. This indicates that the ratio of the first pair of jacknifed eigenvalues has a relatively stable distribution across the levels and types of contamination.

To further examine the performance of the rejection thresholds we calculated the proportion of SWGs which were found to be higher than the threshold, in the various tests. The results for the 18 contaminated tests are summarized in Appendix 5. The proportions are summarized as a function of the eigenvalue examined (first, second or third), the level of contamination and the inter-trait correlation. The overall trend is for the number of unusually large gaps to increase monotonically as a function of the eigenvalue (it is lowest for the first and highest for the third), and the level of contamination, and decrease monotonically with r, the inter-ability correlation. The actual rates of change vary from one threshold to another.

The most important issue, from a practical point of view, is to choose the "best" threshold for detection of wide gaps. To address this issue we focus on the performance of the various indices in the uncontaminated (p=0) case. Table 7 displays the proportion of SWGs exceeding the various indices for the three ratios. Since this is a strictly unidimensional test, we expect this proportion to be invariant for all three ratios and not to exceed its nominal confidence level (95% or 99%). Clearly, 2.25 and the empirical percentiles fail the invariance requirement and the 95% constrained Chebyshev bound is too liberal for the third ratio. In light of these results we conclude that is best to identify as "unusually wide gaps" those values that exceed the 95% unconstrained, or the constrained 99% Chebyshev bounds. We will focus primarily on rejections with 99% confidence. However, for completeness sake, we will report in the sequel results according to all the seven thresholds.

Insert Table 7 about here

Partition of the Tests

Tables 8a - 8c list the maximal SWGs observed in the distributions of the three ratios for each test. The tables also display the pattern of significance achieved by this maximal SWG, and its location. The columns labeled "significance" simply count how many (of the increasingly stringent) thresholds were exceeded in each family of tests. The fixed 2.25 criterion is either surpassed (1 in the table) or not (0). In the 95% and 99% columns, a 1 indicates that the observed value is greater than the empirical percentile but lower than both Chebyshev bounds; a value of 2 describes a situation where the actual value is greater than the constrained (but smaller than the unconstrained) Chebyshev bound, and a value of 3 denotes a case where the maximal gap is larger than the most severe rejection rule. Our previous results (see table 7) dictate to interpret as "significant" values of 2 (at 99%), or values of 3 (at 95%).

The location of the gap is described by reporting the number of items above, and below, it. Recall that according to the logic of RMPA the contaminating items should have lower (i.e. closer to unity) ratios. We rank ordered the ratios in ascending order, so these items are expected to cluster "above" the gap. As a rule, we expect the proportion of item above the gap to match, approximately, the proportion of contamination in the specific test. Since decisions about rejection can be based on the second and/or third eigenvalue, we summarize in table 9 the pattern of results for each test across all three ratios.

We reject the null hypothesis of unidimensionality if:

(1) The number of items "above the gap" < n/2 AND
(2) The Maximal SWG of the second AND/OR the third ratio is greater than the designated rejection threshold.

We examine three rejection rules with decreasing levels of conservatism: (1) 99% according to an unconstrained Chebyshev inequality, (2) 99% according to a constrained Chebyshev inequality, and (3) 95% according to the unconstrained Chebyshev inequality.

Insert Tables 8a - 8c and 9 about here

As expected, there are no significant gaps in the distribution of the first ratio but, in most tests, the largest SWG in the distribution of the second and/or third ratio is significant. We examine these significant gaps according to the three valid rejection thresholds:

The most stringent approach requires a SWG to exceed the 99% threshold derived from a regular Chebyshev inequality. Seven tests have gaps larger than this threshold (three in the distribution of the second ratio, two in the distribution of the third and one in both). Five of these tests have low (p=10) level of contamination, one is moderately (p=25) and the other is highly (p=50) contaminated.

In seven of the remaining tests the Max(SWG) exceeds the 99% threshold derived from a constrained (unimodality + symmetry) Chebyshev inequality. One is uncontaminated (p=0), one is slightly (p=10), three are moderately (25%) and two are highly (p=50) contaminated.

All the other six tests reach significance according to an unconstrained 95% Chebychev bound. This group includes one uncontaminated (p=0) test as well as two moderately (25%) and three highly (p=50) contaminated cases.

All six cases with low (p=10) contamination are significant at the 99% level (five of them by the most severe criterion). In all six cases the gap separates the top 10% items from the bottom 90%. It appears that the procedure works well for this type of contamination.

Only three of the highly contaminated tests (p=50) are significant at 99%. More important, however, is the fact that in all six tests the widest gap is located at the bottom of the distribution. Although the numbers vary slightly across tests, the proportion of items above the gap is always greater than 80%. Clearly, the gap test does not work well for a mixture of two half tests.

The pattern of results is slightly more complex in the case of moderate (p=25) contamination, and it depends on the level of the inter-ability correlation: For both tests with uncorrelated (r=0) abilities, and one of the tests with moderately correlated (r=0.5) abilities, the significant gap (99%) in the distribution of the second ratio separates the upper 25% items from the rest of the test. In the other test with r=0.5 the gap between the top 25% of the items and the lower 75% is significant at the 95% level. Finally, for the tests involving highly correlated abilities (r=0.7), the maximal gap is located at the lower end of the distribution (69 and 72 items above the gap). In both cases the second largest gap distinguishes between the (most) contaminating items and the original ones. Thus, the gap test operates well only for cases with low inter-ability correlations.

To summarize, RMPA found a significant gap in the distribution of the ratios of matched eigenvalues in all the tests examined. In 14 tests the gap was significant at 99% and in the other six at 95%. A significant gap located in the upper half of the distribution (i.e. with fewer items above the gap than bellow it) is taken as a strong indication of violation of unidimensionality and prescribes elimination of all items above the gap. The ten tests identified by this criterion include all those with low contamination (p=10), as well as the moderately contaminated ones (p=25), with moderate level of inter-ability correlation (r <0.7).

In the sequel we focus only on these 10 shortened tests. To facilitate interpretation of the results, we attach plots of the 10 relevant distributions of standardized weighted gaps. The original (1-p)n items are plotted as "*"s and the pn contaminating items are plotted as "C"s.

Note that in all plots:

(i) the contaminating items are clustered at one end of the distribution, and

(ii) there is an unusually large gap separating this cluster from the bulk of the items. This gap can be detected in the raw gaps, but it is more pronounced in the standardized weighted form.

Insert Figures 1-10 about here

The quality of the technique is assessed by its ability to detect the contaminating items and remove them, while retaining the original ones. Table 10 summarizes this analysis for the 10 short tests. For each one we report the hit rate (i.e. contaminating items rejected correctly) and the false alarm rate (i.e. original items rejected incorrectly). The figures are very impressive ---- for all the tests with p=10%, the hit rate is 100% and for the tests with p=25% it is 95%. Both figures are accompanied by false alarm rates close to 0. This impression can be also verified in their ROC curves (e.g. Green & Swets, 1973). These curves plot the hit rate against the false alarm rate for 20 equally spaced rejection thresholds. Each figure includes a curve based on the ratio of the first pair of eigenvalues and one based on the ratio of the second or third (the one that reached significance in that particular test). In all ten cases the latter curve stochastically dominates the former. Furthermore, at practically all points the procedure does a perfect job of detecting the contaminating items.

Insert Table 10 and Figures 11 - 20 about here

<u>Re-examination of the shortened tests</u>

Having shortened 10 tests according to the results of the initial RMPA we repeated steps 4 - 9 of the procedure. The second iteration verifies the unidimensionality of the shortened tests: If the first stage is successful in removing all sources of contamination, we do not expect to detect any significant gaps in this second round.

Tables 11 and 12 report the results of the MPA and the RMPA of the shortened tests. A quick comparison with tables 2 and 4 (summarizing the same results for the original full tests) reveals that all major sources of multidimensionality were eliminated. The ratios of the second eigenvalues, and the ratios of their variances, are close to unity (We assume that a heuristic

MPA would also declare all these tests unidimensional). The third ratios are somewhat higher but are, considerably, lower than those of the original tests.

Insert Tables 11 and 12 about here

The SWGs of the remaining items were calculated, new rejection thresholds were derived, and the gap test was applied again [6]. The results are presented in Tables 13a-13c (parallel in structure and notation to tables 8a-8c).

Insert Tables 13a - 13c about here

As expected, none of the ratios based on the first pair of eigenvalues is significant (according to the 99% Chebyshev bounds). We found significant gaps in the distribution of the second ratio for four tests, and three of them also revealed significant gaps in the third ratio. However with one exception, the significant gaps are in the lower tail of the distribution. Therefore, they are not indicative of violations of unidimensionality. The only exception was the {Rep=B r=0.7 p=10} test. In this case the second iteration of the RMPA prescribes removal of five additional items. All contaminated items were successfully detected by the first iteration so these are five "false alarms". The final test consists of 66 unidimensional items (instead of 72).

SUMMARY

The goal of the current research was to develop a practical, yet theoretically sound and computationally feasible, tool for testing the global dimensionality of large item pools and eliminating items which cause violations of the pool's unidimensionality. Both goals are attained in the unified framework of RMPA.

MPA was developed by Drasgow and Lissak (1983) as an approximate method for testing the unidimensionality of item pools. It relies on a heuristic comparison of a statistic (the second eigenvalue) derived from the matrix of items' intercorrelations and the corresponding value extracted from a "parallel" matrix generated by a unidimensional, and locally independent, model (in our case the three parameter logistic model).

RMPA is based on a similar comparative logic, but improves upon MPA in several ways:

(1) It alleviates some minor technical limitations, through the use of expected correlations under unidimensionality;
(2) It implements a formal test for comparing the observed data set with its parallel (and unidimensional) counterpart.
(3) Contingent upon the results of this test, it provides a method for identifying and eliminating items which violate the test's unidimensionality.

The testing and elimination procedures are based on the "remove one item at a time" principle. This methodology allows one to assess the contribution of each item to the test's eigenvalues. Furthermore, one can determine the variance and distribution of these values and analyze the differential impact of any given item in the observed and parallel matrices. Items which have a "significantly" larger impact in the observed data set violate unidimensionality.

The detection of these items relies on a conservative version of Wainer & Schacht's (1978) "gapping" test. The largest (first) eigenvalues of the observed and expected matrices are practically identical in all cases, regardless of the level of correlation between the two abilities and the degree of contamination. Therefore, we used the distribution of their ratio to determine rejection thresholds for the ratio of the second and third eigenvalues. These thresholds are based on conservative Chebyshev bounds, and are specifically tailored to each test.

RMPA was tested in several simulations of unidimensional item pools which were contaminated by various proportions of items loaded on a secondary nuisance ability. The method was highly successful in identifying low (10%) levels of departure from unidimensionality, and in detecting moderate (25%) deviations from unidimensionality when the abilities were not highly ($r < 0.7$) correlated. In these cases over 90% of the contaminating items were identified and less than 1% of the original items were eliminated erroneously. The procedure failed, and should not be applied, in tests which are equal mixtures (50%) of two abilities.

We conclude by pointing out that the logic of MPA and RMPA can be generalized to other statistics of closeness between the two data sets. For example, it might be interesting to apply it to indices derived from non linear factor analysis (e.s. McDonald, 1982).

REFERENCES

Ackerman, T. (1989) Unidimensional IRT calibration of compensatory and noncompensatory multidimensional items *Applied Psychological Measurement, 13*, 113-127.

Arvesen, J.N. & Salsburg, D.S. (1975) Approximate tests and confidence interva using the jacknife. In R.M. Elashoff (Ed.) *Perspectives in Biometrics*. New York: Academic Press.

Bartlett, M.S. (1950) Tests of significance in factor analysis. *British Journal of Psychology, 3*, 77-85.

Berger, M.P.F. & Knol, D.K. (1990) On the assessment of dimensionality in multidimensional item response theory models. *Paper presented at the annual AERA Meeting*. Boston, Mass.

Cattell, R.B. (1966) The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Cohen, Y. & Bodner, G. (1989) A manual for NITEST: A program for estimating IRT parameters. *Report No. 94. Jerusalem: NITE.*

Cohen, Y., Bodner, G. & Ronen, T. (1989) A manual for NITECAT: A software package for research on IRT/CAT. *Report No. 100. Jerusalem: NITE.*

Collins, L.M., Cliff, N., McCormick, D.J. & Zatlin, J.L. (1986) Factor recovery in binary data sets: A simulation. *Multivariate Behavioral Research, 23*, 377-392.

Crawford, C.B. & Koopman, P. (1973) A note on Horn's test for the number of factors in factor analysis. *Multivariate Behavioral Research, 8*, 117-125.

Devlin, S.J., Gnanadesikan, R., & Kettenring, J.R. (1975) Robust estimation and outlier detection with correlation coefficients. *Biometrika, 62*, 531-545.

Drasgow, F. & Lissak, R.I. (1983) Modified parallel analysis: A procedure for examining the latent dimensionality of dichotomously scored item responses. *Journal of Applied Psychology, 68*, 363 - 373.

Drasgow, F. & Parsons, C. (1983) Applications of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement, 7*, 189-199.

Dorans, N.J. & Kingston, N.M. (1985) The effects of violations of unidimensionality on the estimation of item and ability parameters and on item response theory equating to the GRE verbal scale. *Journal of Educational Measurement, 22*, 249-262.

Green, B.F. (1983) The promise of tailored tests. In H. Wainer and S. Messick (Eds.) *Principals of Modern Psychological Measurement*. Hillsdale, NJ: Lawrence Earlbaum Associates Press.

Green, D.M. & Swets, J.A. (1973) *Signal detection theory and psychophysics.* New York: Robert E. Krieger Publishing Co.

Hambleton, R.K. (1983) *Applications of Item Response Theory.* Vancouver BC: Educational Research Institute of British Columbia.

Hampel, F.R. (1974) The influence curve and its role in robust estimation. *Journal of the American Statistical Association, 69,* 383-393.

Hattie, J. (1984) An empirical study of various indices for determining unidimensionality. *Multivariate Behavioral Research, 19,* 49-78.

Hattie, J. (1985) Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement, 9,* 139-164.

Horn, J.L. (1965) A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179-164.

Hulin, C.L., Drasgow, F. & Parsons, C.K. (1983) *Item Response Theory.* Homewood, Ill: Dow Jones-Irwin Publishing Company.

Humphreys, L.G. (1985) General intelligence: An integration of factor, test and simplex theory. IN B.B. Wolman (Ed.) *Handbook of Intelligence* (pp. 201-224). New York: Wiley.

Humphreys, L.G. & Montanelli, R.G. (1975) An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research, 10,* 193-205.

Kaiser, H.F. (1960) The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151.

Kendall, M. & Stuart, A. (1979) *The Advanced Theory of Statistics (Vol 2),* London: McMillan, Fourth Edition.

Knol, D.L. & Berger, M.P.F. (1991) Empirical comparison between factor analysis and multidimensional item response models. *Multivariate Behavioral Research, 26,* 457-478.

Longman, R.S., Cota, A.A., Holden, R.R. & Fekken, G.C. (1989) A regression equation for the parallel analysis criterion in principal components analysis: Mean and 95th percentile eigenvalues. *Multivariate Behavioral Research, 24,* 59-69.

Lord, F.M. (1980) *Applications of Item Response Theory to Practical Testing Problems.* Hillsdale, NJ: Lawrence Erlbaum Publishers.

Lord, F.M., & Novick, M. R. (1968) *Statistical Theories of Mental Test Scores.* Reading,MA: Addison-Wesley Publishing Company.

McDonald, R.P. (1981) The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34,* 100-117.

McDonald, R.P. (1982) Linear vs. nonlinear models in item response theory. *Applied Psychological Measurement, 6,* 379-396.

Miller, R.G. (1974) The jacknife - A review. *Biometrika, 61,* 1-15.

Miller, M.D, & Oshima, T.C. (1992) Effect of sample size, number of biased items, and magnitude of bias on a two-stage item bias estimation method. *Applied Psychological Measurement, 16,* 381-388.

Mosteller, F. & Tukey, J.W. (1977) *Data Analysis and Regression: A Second Course in Statistics.* Reading, Mass: Addison-Wesley Publishing Co.

Mulaik, S.A. (1972) *The Foundations of Factor Analysis.* New York:McGraw Hill Book Company.

Nandakumar, R. (1991) Traditional dimensionality versus essential dimensionality. *Journal of Educational Measurement, 28,* 99-117.

Oshima, T.C. & MIller, M.D. (1992) Multidimensionality and item bias in item response theory. *Applied Psychological Measurement, 16,* 237-248.

Reckase, M.D. (1979) Unifactor latent trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics, 4,* 207-230..

Reckase, M.D. (1985) The difficulty of test items that measure more than one ability. *Applied Psychological Measurement, 9,* 401-412.

Reckase, M.D., Ackerman, T.A., & Carlson, J.E. (1988) Building unidimensional tests using multidimensional items. *Journal of Educational Measurement, 25,* 193-203.

Roznowsky, M., Tucker, L.R., & Humphreys, L.G. (1991) Three approaches to determining the dimensionality of binary items. *Applied Psychological Measurement, 15,* 109-127.

Saw, J.G., Yang, M.C.K, & Mo, T.C. (1984) Chebyshev inequality with estimated mean and variance. *The American Statistician, 38,* 130-132.

Stout, W.F. (1987) A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika, 52,* 589-617.

Stout, W.F. (1990) A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika, 55,* 293-325.

Stuart, A., & Ord, J.K. (1987) *Kendall's Advanced Theory of Statistics (Vol 1)*, New York: Oxford University Press, Fifth Edition.

Sympson, J.B. (1978) A model for testing with multidimensional items. In Weiss, D.J. (Ed.) *Proceedings of the 1977 CAT Conference* (pp 82-98). Minneapolis: University of Minnesota, Department of Psychology.

Traub, R.E. (1983) A priori considerations in choosing an item response model. In R.K. Hambleton (Ed.) *Applications of Item Response Theory* (pp. 55-70). Vancouver BC:Educational Research Institute of British Columbia.

Wainer, H. & Schacht, S. (1978) Gapping. *Psychometrika, 43,* 203-212.

Weiss, D.J. (Ed.) (1983) *New Horizons in Testing: Latent Trait Test Theory and Computerized Adaptive Testing.* New York: Academic Press.

Yen, W. M. (1984) Effects of local item dependence on the fit and equating performance of the three parameter logistic model. *Applied Psychological Measurement, 8,* 125-145.

Yen, W. M. (1985) Increasing item complexity: A possible cause of scale shrinkage for unidimensional item response theory. *Psychometrika, 50,* 399-410.

Zwick, W.R. & Velicer, W.F. (1986) Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432-442.

FOOTNOTES

(1) Strictly speaking "jacknifing" refers to an analysis in which observations (i.e. respondents) are eliminated one at a time from the sample. In this case, we eliminate variables (items) in a similar fashion. Several item analysis computer programs use a similar approach in order to identify subscales with maximal reliability.

(2) I_i, the sample influence function (Devlin, Gnanadesikan and Kettenring, 1975; Hampel, 1974) of parameter, T, is given by:

$$I_i = (n-1)(T - T_{-i}),$$

where n is the number of items, and T_{-i} is an estimate of the parameter T obtained after the elimination of item i. Note that I_i is, simply, a linear transformation of T_{-i}.

(3) The term "conservative" is used here according to the standard convention in statistical inference, i.e. a procedure is more conservative than its competitor if it invokes a more stringent criterion in rejecting the null hypothesis.

(4) Strictly speaking, Chebyshev inequality requires knowledge of the parameters (mean and variance) *of the population of interest.* However, Saw, Yang and Mo (1984) have shown that sample estimates of these parameters can be used, with very little loss of precision, in moderately large samples.

(5) Occasionally a large (and significant) gap will be detected in the lower tail of the distribution, i.e. separating the bulk of the data from a minority of items with unusual low ratios of observed/expected jacknifed eigenvalues. Clearly, these cases are not relevant for our goal.

(6) Since the procedure is data driven, we opt not to use the thresholds values employed in the first stage. Thus, when analyzing a test consisting of $(n - m_1)$ items one should obtain the same results, and reach the same conclusions, whether it is treated as "an original" or "a reduced" test.

Table 1:

Means, standard deviations and correlations of the two sets of item parameters

Rep=B Parameter	n	Mean	Std. Dev.		Rep=R Parameter	n	Mean	Std. Dev.
a	80	1.123	0.245		a	80	1.328	0.511
b	80	0.172	0.873		b	80	-0.026	0.985
c	80	0.202	0.057		c	80	0.161	0.098
		Correlations					Correlations	
	a	b	c			a	b	c
a	1.000	0.090	-0.196		a	1.000	0.518	0.397
b	0.090	1.000	-0.260		b	0.518	1.000	0.754
c	-0.196	-0.260	1.000		c	0.397	0.754	1.000

Table 2:

Modified Parallel Analysis (MPA) of 20 tests: The first three eigenvalues for the observed and expected matrices, and their ratios

			Eigenvalue 1			Eigenvalue 2			Eigenvalue 3		
Rep	r	p	Exp	Obs	Obs/Exp	Exp	Obs	Obs/Exp	Exp	Obs	Obs/Exp
B	-	0	24.34	25.10	1.03	1.79	1.78	0.99	0.17	0.67	4.02
B	0.0	10	22.03	22.71	1.03	1.62	2.55	1.58	0.17	1.66	9.72
B	0.0	25	18.15	18.80	1.04	1.43	5.59	3.92	0.15	1.53	10.17
B	0.0	50	11.93	12.58	1.05	0.61	12.01	19.77	0.08	0.85	10.35
B	0.5	10	22.72	23.55	1.04	1.66	1.78	1.07	0.18	1.66	9.13
B	0.5	25	12.93	20.77	1.04	1.46	3.87	2.65	0.16	1.54	9.75
B	0.5	50	17.76	18.71	1.05	0.92	6.05	6.60	0.07	1.07	14.38
B	0.7	10	23.31	24.16	1.04	1.66	1.73	1.04	0.18	1.26	7.09
B	0.7	25	21.27	22.13	1.04	1.46	2.30	1.58	0.15	1.55	10.51
B	0.7	50	19.91	20.82	1.05	1.18	3.67	3.11	0.09	1.34	14.26
R	-	0	26.20	26.23	1.00	3.47	3.22	0.93	0.36	0.63	1.78
R	0.0	10	23.59	23.84	1.01	2.84	2.86	1.01	0.25	2.57	10.20
R	0.0	25	19.57	19.81	1.01	2.38	6.20	2.60	0.20	2.33	11.91
R	0.0	50	12.23	12.90	1.05	1.45	12.20	8.39	0.15	1.77	11.53
R	0.5	10	23.68	24.14	1.02	2.76	2.76	1.00	0.28	1.90	6.69
R	0.5	25	21.45	21.90	1.02	2.48	4.31	1.74	0.25	2.45	9.70
R	0.5	50	19.30	19.88	1.03	1.79	6.35	3.55	0.13	1.93	14.89
R	0.7	10	24.45	24.86	1.02	2.88	2.87	1.00	0.30	1.18	4.01
R	0.7	25	22.91	23.30	1.02	2.64	3.16	1.20	0.23	2.28	10.09
R	0.7	50	21.78	22.25	1.02	2.32	3.89	1.68	0.17	2.39	14.17

Notes:
All results based on n=80 items and N=2000 respondents.
Exp = Derived from matrix of expected correlations
Obs = Derived from matrix of observed correlations.

Table 3:

Revised Modified Parallel Analysis (RMPA) of 20 tests: Means and standard deviations of the first three eigenvalues of the jacknifed submatrices (observed and expected)

Rep	r	p	Source	Eigenvalue 1 Mean	SD	Eigenvalue 2 Mean	SD	Eigenvalue 3 Mean	SD
B	-	0	Exp	24.036	0.122	1.764	0.029	0.165	0.004
			Obs	24.786	0.118	1.756	0.027	0.664	0.009
B	0.0	10	Exp	21.752	0.145	1.598	0.026	0.168	0.004
			Obs	22.429	0.147	2.515	0.103	1.637	0.027
B	0.0	25	Exp	17.920	0.163	1.409	0.027	0.148	0.004
			Obs	18.563	0.169	5.519	0.143	1.507	0.027
B	0.0	50	Exp	11.778	0.136	0.600	0.013	0.081	0.002
			Obs	12.432	0.158	11.842	0.165	0.842	0.016
B	0.5	10	Exp	22.437	0.133	1.640	0.027	0.179	0.004
			Obs	23.255	0.128	1.769	0.036	1.630	0.038
B	0.5	25	Exp	19.685	0.136	1.443	0.026	0.156	0.004
			Obs	20.507	0.128	3.818	0.084	1.525	0.027
B	0.5	50	Exp	17.536	0.089	0.906	0.014	0.073	0.002
			Obs	18.472	0.085	5.972	0.035	1.053	0.019
B	0.7	10	Exp	23.023	0.125	1.640	0.026	0.175	0.004
			Obs	23.858	0.119	1.713	0.025	1.241	0.043
B	0.7	25	Exp	21.008	0.126	1.437	0.024	0.145	0.003
			Obs	21.850	0.122	2.267	0.048	1.526	0.027
B	0.7	50	Exp	19.660	0.103	1.164	0.019	0.093	0.002
			Obs	20.564	0.101	3.625	0.025	1.326	0.021
R	-	0	Exp	25.874	0.128	3.424	0.040	0.351	0.006
			Obs	25.899	0.132	3.178	0.036	0.628	0.007
R	0.0	10	Exp	23.292	0.155	2.808	0.040	0.249	0.006
			Obs	23.547	0.160	2.829	0.035	2.531	0.113
R	0.0	25	Exp	19.330	0.174	2.354	0.039	0.193	0.005
			Obs	19.561	0.182	6.121	0.149	2.301	0.035
R	0.0	50	Exp	12.078	0.142	1.436	0.031	0.152	0.004
			Obs	12.735	0.187	12.045	0.175	1.751	0.037
R	0.5	10	Exp	23.389	0.141	2.726	0.038	0.281	0.006
			Obs	23.835	0.138	2.726	0.034	1.876	0.078
R	0.5	25	Exp	21.181	0.152	2.449	0.038	0.249	0.006
			Obs	21.629	0.143	4.251	0.084	2.415	0.035
R	0.5	50	Exp	19.063	0.107	1.763	0.023	0.128	0.004
			Obs	19.636	0.101	6.263	0.037	1.906	0.025
R	0.7	10	Exp	24.145	0.136	2.839	0.038	0.291	0.006
			Obs	24.552	0.132	2.835	0.033	1.166	0.043
R	0.7	25	Exp	22.627	0.138	2.605	0.037	0.223	0.005
			Obs	23.006	0.132	3.123	0.040	2.252	0.033
R	0.7	50	Exp	21.504	0.118	2.292	0.030	0.166	0.004
			Obs	21.970	0.116	3.843	0.027	2.357	0.030

Notes:
All results based on n=80 items and N=2000 respondents.
Exp = Derived from matrix of expected correlations
Obs = Derived from matrix of observed correlations.

Table 4:

Revised Modified Parallel Analysis (RMPA) of 20 tests:
Ratio of means and variances of eigenvalues of the jacknifed submatrices
(Ratio = observed / expected)

Rep	r	p	Eigenvalue 1 Mean	Eigenvalue 1 Var	Eigenvalue 2 Mean	Eigenvalue 2 Var	Eigenvalue 3 Mean	Eigenvalue 3 Var
B	-	0	1.031	0.933	0.995	0.884	4.030	6.007
B	0.0	10	1.031	1.034	1.574	15.449	9.730	44.102
B	0.0	25	1.036	1.074	3.917	28.726	10.187	52.047
B	0.0	50	1.056	1.355	19.748	151.159	10.350	83.432
B	0.5	10	1.036	0.920	1.078	1.779	9.084	80.253
B	0.5	25	1.042	0.885	2.646	10.454	9.763	44.765
B	0.5	50	1.053	0.909	6.594	5.886	14.379	133.295
B	0.7	10	1.036	0.916	1.045	0.908	7.086	116.778
B	0.7	25	1.040	0.934	1.578	3.836	10.524	62.910
B	0.7	50	1.046	0.955	3.114	1.681	14.278	92.026
R	-	0	1.001	1.054	0.928	0.804	1.790	1.296
R	0.0	10	1.011	1.067	1.008	0.776	10.183	370.061
R	0.0	25	1.012	1.094	2.600	14.554	11.918	53.081
R	0.0	50	1.054	1.746	8.387	31.775	11.527	78.033
R	0.5	10	1.019	0.946	1.000	0.786	6.685	165.024
R	0.5	25	1.021	0.888	1.736	4.773	9.705	36.368
R	0.5	50	1.030	0.892	3.554	2.554	14.892	45.144
R	0.7	10	1.017	0.944	0.999	0.770	4.004	44.388
R	0.7	25	1.017	0.916	1.199	1.205	10.082	42.093
R	0.7	50	1.022	0.966	1.677	0.763	14.166	68.059

Notes:
All results based on n=80 items and N=2000 respondents.

Table 5:

Revised Modified Parallel Analysis (RMPA) of 20 tests:
Correlations of eigenvalues of the observed and the expected jacknifed submatrices

Rep	r	p	Ev 1	Ev 2	Ev 3		Rep	r	p	Ev 1	Ev 2	Ev 3
B	-	0	0.996	0.888	0.650		R	-	0	0.976	0.956	0.195
B	0.0	10	0.995	-0.237	0.603		R	0.0	10	0.995	0.963	-0.147
B	0.0	25	0.996	-0.337	0.655		R	0.0	25	0.996	-0.414	0.306
B	0.0	50	0.981	-0.459	-0.049		R	0.0	50	-0.641	0.537	0.321
B	0.5	10	0.997	-0.256	0.299		R	0.5	10	0.998	0.968	-0.129
B	0.5	25	0.998	-0.320	0.593		R	0.5	25	0.996	-0.320	0.394
B	0.5	50	0.990	-0.180	0.310		R	0.5	50	0.996	-0.007	0.218
B	0.7	10	0.997	0.863	-0.111		R	0.7	10	0.998	0.968	-0.111
B	0.7	25	0.996	-0.311	0.608		R	0.7	25	0.995	0.350	0.140
B	0.7	50	0.997	-0.206	0.508		R	0.7	50	0.996	0.060	0.127

Notes:
All results based on n=80 items and N=2000 respondents.
Ev = Eigenvalue

Table 6:

Revised Modified Parallel Analysis (RMPA) of 20 tests:
Seven detection thresholds based on the distribution of the standardized Weighted Gaps (SWGs) based on the ratio of the first observed and expected jacknifed eigenvalues

					Threshold						
			SWG			95 %			99 %		
Rep	r	p	Mean	S.D.	2.25	Emp	UChe	Cheb	Emp	UChe	Cheb
B	-	0	0.94	0.49	2.25	1.89	2.42	3.17	2.38	4.24	5.88
B	0.0	10	0.91	0.58	2.25	1.94	2.67	3.54	2.58	4.81	6.76
B	0.0	25	0.97	0.52	2.25	1.96	2.54	3.32	2.87	4.45	6.19
B	0.0	50	0.84	0.56	2.25	2.07	2.53	3.37	3.09	4.59	6.46
B	0.5	10	0.90	0.42	2.25	1.61	2.15	2.77	1.92	3.67	5.06
B	0.5	25	0.96	0.53	2.25	2.01	2.56	3.36	2.93	4.52	6.30
B	0.5	50	0.97	0.51	2.25	1.84	2.50	3.26	2.26	4.37	6.07
B	0.7	10	0.98	0.50	2.25	1.89	2.48	3.23	2.80	4.31	5.97
B	0.7	25	0.95	0.57	2.25	2.13	2.65	3.50	2.66	4.73	6.61
B	0.7	50	0.89	0.52	2.25	1.92	2.45	3.23	2.13	4.35	6.08
B	Mean		0.93	0.52	2.25	1.93	2.49	3.27	2.56	4.40	6.14
R	-	0	0.95	0.44	2.25	1.65	2.27	2.93	2.12	3.89	5.36
R	0.0	10	0.99	0.56	2.25	1.85	2.66	3.50	3.40	4.72	6.58
R	0.0	25	0.94	0.54	2.25	1.90	2.56	3.37	2.65	4.54	6.33
R	0.0	50	0.79	0.51	2.25	1.77	2.30	3.06	2.47	4.16	5.84
R	0.5	10	0.99	0.48	2.25	1.82	2.42	3.14	1.93	4.18	5.78
R	0.5	25	0.95	0.52	2.25	2.03	2.52	3.31	2.24	4.44	6.18
R	0.5	50	0.93	0.53	2.25	1.82	2.53	3.33	3.00	4.48	6.25
R	0.7	10	1.07	0.61	2.25	2.16	2.90	3.81	2.87	5.12	7.14
R	0.7	25	0.92	0.45	2.25	1.95	2.28	2.95	2.23	3.93	5.43
R	0.7	50	0.96	0.47	2.25	1.85	2.38	3.09	2.03	4.12	5.70
R	Mean		0.95	0.51	2.25	1.88	2.48	3.25	2.49	4.35	6.06
Mean			0.94	0.52	2.25	1.90	2.49	3.26	2.53	4.38	6.10

Notes:
All results based on n=80 items and N=2000 respondents.
Emp = Empirical distribution
UChe = Chebyshev bound assuming unimodality
Cheb = Chebyshev bound

Table 7:

Revised Modified Parallel Analysis (RMPA):
Proportion of Standardized Weighted Gaps (SWGs) exceeding each of the seven thresholds in the uncontaminated unidimensional test

		Threshold					
		95 %				99 %	
Eigenvalue	2.250	Emp	UChe	Cheb	Emp	UChe	Cheb
1	0.006	0.051	0.000	0.000	0.013	0.000	0.000
2	0.082	0.177	0.063	0.019	0.070	0.006	0.000
3	0.120	0.215	0.108	0.038	0.120	0.006	0.000
Mean	0.069	0.148	0.057	0.019	0.068	0.004	0.000

Notes:
All results based on n=80 items and N=2000 respondents.
Emp = Empirical distribution
UChe = Chebyshev bound assuming unimodality
Cheb = Chebyshev bound

Table 8a:

Revised Modified Parallel Analysis (RMPA) of 20 tests: Maximal Standardized Weighted Gap (SWG) and significance according to three types of thresholds (First eigenvalue)

						Significance*		No. of items	
Rep	r	p	Gap	Max(SWG)	2.25	95%	99%	Below	Above
B	-	0	0.00007800	2.37665	1	1	1	40	40
B	0.0	10	0.00011536	2.58303	1	1	1	29	51
B	0.0	25	0.00017863	2.87133	1	2	1	52	28
B	0.0	50	0.00074733	3.09437	1	2	1	46	34
B	0.5	10	0.00007026	1.92493	0	1	1	42	38
B	0.5	25	0.00023166	2.93434	1	2	1	20	60
B	0.5	50	0.00011191	2.25905	1	1	1	49	31
B	0.7	10	0.00011347	2.80103	1	2	1	35	45
B	0.7	25	0.00059257	2.66357	1	2	1	76	4
B	0.7	50	0.00035381	2.12819	0	1	1	76	4
R	-	0	0.00013665	2.12450	0	1	1	30	50
R	0.0	10	0.00083803	3.39892	1	2	1	4	76
R	0.0	25	0.00019422	2.65083	1	2	1	22	58
R	0.0	50	0.00513373	2.47295	1	2	1	40	40
R	0.5	10	0.00004417	1.92817	0	1	1	33	47
R	0.5	25	0.00010552	2.24198	0	1	1	38	42
R	0.5	50	0.00017510	2.99795	1	2	1	28	52
R	0.7	10	0.00010866	2.87002	1	1	1	16	64
R	0.7	25	0.00009185	2.22529	0	1	1	33	47
R	0.7	50	0.00017596	2.02883	0	1	1	7	73

**Note:*

1 --> Max(SWG) > Empirical percentile

2 --> Max(SWG) > Chebyshev + unimodality

3 --> Max(SWG) > Chebyshev

Table 8b:

Revised Modified Parallel Analysis (RMPA) of 20 tests:
Maximal Standardized Weighted Gap (SWG) and significance according to three types of thresholds (Second eigenvalue)

						Significance*		No. of items	
Rep	r	p	Gap	Max(SWG)	2.25	95%	99%	Below	Above
B	-	0	0.002948	4.6028	1	3	2	13	67
B	0.0	10	0.154001	10.3842	1	3	3	72	8
B	0.0	25	0.058056	5.5693	1	3	2	61	19
B	0.0	50	0.420034	2.9129	1	2	0	6	74
B	0.5	10	0.061734	7.9202	1	3	3	72	8
B	0.5	25	0.020733	3.9763	1	3	1	62	18
B	0.5	50	0.016864	2.7078	1	2	1	30	50
B	0.7	10	0.017799	3.6832	1	3	1	78	2
B	0.7	25	0.017849	4.2707	1	3	1	64	16
B	0.7	50	0.061186	2.6536	1	2	1	3	77
R	-	0	0.002074	2.6308	1	2	1	4	76
R	0.0	10	0.001358	3.1794	1	2	0	10	70
R	0.0	25	0.041428	5.1730	1	3	2	60	20
R	0.0	50	0.063916	7.3824	1	3	3	15	65
R	0.5	10	0.002657	2.7704	1	2	1	4	76
R	0.5	25	0.032784	6.7929	1	3	3	61	19
R	0.5	50	0.027717	3.0870	1	2	1	7	73
R	0.7	10	0.000493	2.7074	1	1	0	23	57
R	0.7	25	0.005675	2.7842	1	2	1	71	9
R	0.7	50	0.005415	2.7835	1	2	1	19	61

**Note:*
1 --> Max(SWG) > Empirical percentile
2 --> Max(SWG) > Chebyshev + unimodality
3 --> Max(SWG) > Chebyshev

Table 8c:

Revised Modified Parallel Analysis (RMPA) of 20 tests:
Maximal Standardized Weighted Gap (SWG) and significance according to three types of thresholds (Third eigenvalue)

					Significance*			No. of items	
Rep	r	p	Gap	Max(SWG)	2.25	95%	99%	Below	Above
B	-	0	0.197155	5.0647	1	3	2	2	78
B	0.0	10	0.407039	3.5361	1	2	1	1	79
B	0.0	25	0.087976	4.3483	1	3	1	9	71
B	0.0	50	0.065372	3.5205	1	3	1	12	68
B	0.5	10	0.110290	3.4420	1	3	1	73	7
B	0.5	25	0.138565	4.4914	1	3	1	6	74
B	0.5	50	0.237589	5.5091	1	3	2	9	71
B	0.7	10	0.184939	7.1549	1	3	3	71	9
B	0.7	25	0.140215	5.1461	1	3	2	8	72
B	0.7	50	0.116452	3.3573	1	3	1	7	73
R	-	0	0.013257	3.6044	1	3	1	9	71
R	0.0	10	0.419092	9.4241	1	3	3	72	8
R	0.0	25	0.115029	4.1520	1	3	1	10	70
R	0.0	50	0.225673	8.3301	1	3	3	10	70
R	0.5	10	0.287058	10.3919	1	3	3	72	8
R	0.5	25	0.118107	5.3490	1	3	2	12	68
R	0.5	50	0.464085	4.3802	1	3	1	3	77
R	0.7	10	0.111553	6.4281	1	3	2	72	8
R	0.7	25	0.072746	3.5625	1	3	1	11	69
R	0.7	50	0.272261	4.7821	1	3	2	7	73

**Note:*

1 --> Max(SWG) > Empirical percentile
2 --> Max(SWG) > Chebyshev + unimodality
3 --> Max(SWG) > Chebyshev

Table 9:

Revised Modified Parallel Analysis (RMPA) of 20 tests: Maximal Standardized Weighted Gaps (SWGs) and significance according to all eigenvalues

			Max (SWG)			2.25+	95%*	99%*
Rep	r	p	Z1	Z2	Z3	1 2 3	1 2 3	1 2 3
B	-	0	2.38	4.60	5.06	1 1 1	1 3 3	1 2 2
B	0.0	10	2.58	10.38	3.54	1 1 1	1 3 2	1 3 1
B	0.0	25	2.87	5.57	4.35	1 1 1	2 3 3	1 2 1
B	0.0	50	3.09	2.91	3.52	1 1 1	2 2 3	1 0 1
B	0.5	10	1.92	7.92	3.44	0 1 1	1 3 3	1 3 1
B	0.5	25	2.93	3.98	4.49	1 1 1	2 3 3	1 1 1
B	0.5	50	2.26	2.71	5.51	1 1 1	1 2 3	1 1 2
B	0.7	10	2.80	3.68	7.15	1 1 1	2 3 3	1 1 3
B	0.7	25	2.66	4.27	5.15	1 1 1	2 3 3	1 1 2
B	0.7	50	2.13	2.65	3.36	0 1 1	1 2 3	1 1 1
R	-	0	2.12	2.63	3.60	0 1 1	1 2 3	1 1 1
R	0.0	10	3.40	3.18	9.42	1 1 1	2 2 3	1 0 3
R	0.0	25	2.65	5.17	4.15	1 1 1	2 3 3	1 2 1
R	0.0	50	2.47	7.38	8.33	1 1 1	2 3 3	1 3 3
R	0.5	10	1.93	2.77	10.39	0 1 1	1 2 3	1 1 3
R	0.5	25	2.24	6.79	5.35	0 1 1	1 3 3	1 3 2
R	0.5	50	3.00	3.09	4.38	1 1 1	2 2 3	1 1 1
R	0.7	10	2.87	2.71	6.60	1 1 1	1 1 3	1 0 2
R	0.7	25	2.23	2.78	3.56	0 1 1	1 2 3	1 1 1
R	0.7	50	2.03	2.78	4.78	0 1 1	1 2 3	1 1 2

Note: *

1 --> Max(SWG) > Empirical percentile

2 --> Max(SWG) > Chebyshev + unimodality

3 --> Max(SWG) > Chebyshev

+ 1 -- > Max(SWG) > 2.25

Table 10:

Revised Modified Parallel Analysis (RMPA) of 10 short tests:
Total number of items eliminated and accuracy of the elimination procedure

			Items eliminated			
Rep	r	p	Total	% of "hits"	% of "false alarms"	Significant Eigenvalue
B	0.0	10	8	100	0	2
B	0.5	10	8	100	0	2
B	0.7	10	9	100	1	3
R	0.0	10	8	100	0	3
R	0.5	10	8	100	0	3
R	0.7	10	8	100	0	3
Mean			8.2	100	0.2	-
B	0.0	25	19	95	0	2
*B	0.5	25	18	90	0	2
R	0.0	25	20	100	0	2
R	0.5	25	19	95	0	2
Mean			19	95	0	
Mean			-	98	0.1	

Note:
Tests shortened by 99% criterion
* These tests shortened by a 95% criterion

Table 11:

Modified Parallel Analysis (MPA) of 10 short tests:
The first three eigenvalues for the observed and expected matrices, and their ratios

			Eigenvalue 1			Eigenvalue 2			Eigenvalue 3		
Rep	r	p	Exp	Obs	Obs/Exp	Exp	Obs	Obs/Exp	Exp	Obs	Obs/Exp
B	0.0	10	21.86	22.71	1.04	1.62	1.65	1.02	0.17	0.62	3.64
B	0.0	25	17.89	18.78	1.05	1.42	1.48	1.05	0.14	0.56	3.86
B	0.5	10	21.95	22.71	1.03	1.65	1.65	1.00	0.18	0.62	3.51
B	0.5	25	18.00	18.85	1.05	1.40	1.49	1.06	0.15	0.58	3.89
B	0.7	10	21.49	22.27	1.04	1.58	1.61	1.01	0.17	0.62	3.63
R	0.0	10	23.35	23.84	1.02	2.84	2.85	1.01	0.25	0.54	2.19
R	0.0	25	19.17	19.79	1.03	2.37	2.30	0.98	0.19	0.50	2.63
R	0.5	10	22.79	23.21	1.02	2.70	2.71	1.00	0.28	0.58	2.10
R	0.5	25	19.27	19.73	1.02	2.33	2.45	1.05	0.24	0.53	2.21
R	0.7	10	22.83	23.21	1.02	2.73	2.71	0.99	0.28	0.58	2.04

Notes:
All results based on N=2000 respondents, and various number of items.
Exp = Derived from matrix of expected correlations
Obs = Derived from matrix of observed correlations.

Table 12:

Revised Modified Parallel Analysis (RMPA) of 10 short tests:
Ratio of means and variances of eigenvalues of the jacknifed submatrices
(Ratio = observed / expected)

			Eigenvalue 1		Eigenvalue 2		Eigenvalue 3	
Rep	r	p	Mean	Var	Mean	Var	Mean	Var
B	0.0	10	1.039	0.928	1.023	1.066	3.660	5.473
B	0.0	25	1.049	0.923	1.046	1.032	3.889	4.006
B	0.5	10	1.035	0.922	1.005	0.983	3.526	5.112
B	0.5	25	1.047	0.959	1.064	1.065	3.913	6.538
B	0.7	10	1.036	0.919	1.015	1.072	3.641	5.785
R	0.0	10	1.021	0.975	1.005	0.816	2.202	1.152
R	0.0	25	1.033	0.947	0.976	0.760	2.648	1.670
R	0.5	10	1.018	0.960	1.001	0.785	2.107	1.299
R	0.5	25	1.024	0.997	1.054	0.863	2.227	1.418
R	0.7	10	1.017	0.949	0.991	0.772	2.047	1.102

Table 13a:

Revised Modified Parallel Analysis (RMPA) of 10 short tests: Maximal Standardized Weighted Gap (SWG) and significance according to three types of thresholds First Eigenvalue

					Significance*			No. of items		
Rep	r	p	Gap	Max(SWG)	2.25	95%	99%	Below	Above	Total
B	0.0	10	0.00012884	4.38366	1	1	1	21	51	72
B	0.0	25	0.00017634	4.77832	1	2	1	33	28	61
B	0.5	10	0.00011606	4.63637	1	1	1	45	27	72
B	0.5	25	0.00017343	1.01824	0	1	1	10	52	62
B	0.7	10	0.00008747	2.41954	1	1	1	26	45	71
R	0.0	10	0.00005516	3.73614	1	1	1	34	38	72
R	0.0	25	0.00025199	0.44202	0	1	1	8	52	60
R	0.5	10	0.00006304	3.83447	1	1	1	32	40	72
R	0.5	25	0.00010094	2.84225	1	1	1	39	22	61
R	0.7	10	0.00006237	0.75571	0	1	1	32	40	72

**Note:*

1 --> Max(SWG) > Empirical percentile
2 --> Max(SWG) > Chebyshev + unimodality
3 --> Max(SWG) > Chebyshev

Table 13b:

Revised Modified Parallel Analysis (RMPA) of 10 short tests:
Maximal Standardized Weighted Gap (SWG) and significance according to three types of thresholds.
Second Eigenvalue

Rep	r	p	Gap	Max(SWG)	Significance*			No. of items		
					2.25	95%	99%	Below	Above	Total
B	0.0	10	0.0012185	6.23878	1	3	1	20	52	72
B	0.0	25	0.0025284	6.06745	1	2	1	10	51	61
B	0.5	10	0.0062079	5.52540	1	2	1	69	3	72
B	0.5	25	0.0241364	5.76306	1	3	3	1	61	62
B	0.7	10	0.0053278	6.40449	1	3	2	66	5	71
R	0.0	10	0.0007408	0.16707	0	0	0	21	51	72
R	0.0	25	0.0023976	1.68404	0	3	3	5	55	60
R	0.5	10	0.0059670	4.56087	1	2	1	2	70	72
R	0.5	25	0.0039343	3.44740	1	2	1	4	57	61
R	0.7	10	0.0012628	4.00428	1	3	3	9	63	72

**Note:*
1 --> Max(SWG) > Empirical percentile
2 --> Max(SWG) > Chebyshev + unimodality
3 --> Max(SWG) > Chebyshev

Table 13c:

Revised Modified Parallel Analysis (RMPA) of 10 short tests:
Maximal Standardized Weighted Gap (SWG) and significance according to three types of thresholds
Third Eigenvalue

					Significance*			No. of items		
Rep	r	p	Gap	Max(SWG)	2.25	95%	99%	Below	Above	Total
B	0.0	10	0.056692	4.94463	1	2	1	4	68	72
B	0.0	25	0.170522	5.36876	1	2	1	3	58	61
B	0.5	10	0.015446	4.50512	1	1	0	14	58	72
B	0.5	25	0.240886	4.33449	1	3	3	1	61	62
B	0.7	10	0.017104	3.98031	1	3	1	14	57	71
R	0.0	10	0.045500	4.07539	1	1	1	5	67	72
R	0.0	25	0.045556	4.54326	1	3	3	10	50	60
R	0.5	10	0.015318	1.00964	0	0	0	14	58	72
R	0.5	25	0.021710	0.80738	0	0	0	14	47	61
R	0.7	10	0.022073	1.57497	0	3	2	15	57	72

**Note:*
1 --> Max(SWG) > Empirical percentile
2 --> Max(SWG) > Chebyshev + unimodality
3 --> Max(SWG) > Chebyshev

FIGURE CAPTIONS

Figure 18: ROC curves for the ratio of the first and second pair of eigenvalues (Rep=B, r=0.5, p=25).

Figure 19: ROC curves for the ratio of the first and second pair of eigenvalues (Rep=R, r=0, p=25).

Figure 20: ROC curves for the ratio of the first and second pair of eigenvalues (Rep=R, r=0.5, p=25).

Figure 1: **Distribution of SWGs based on the ratios of the second pair of eigenvalues (REP=B R=0 P=10)**

	RL_2	SWG_2
1	1.32	0.00
2	1.34	1.43
3	1.35	1.50
4	1.36	0.63
5	1.36	0.88
6	1.39	3.53
7	1.40	3.05
8	1.42	3.23
9	1.57	10.38
10	1.57	1.11
11	1.58	0.34
12	1.58	0.67
13	1.58	0.41
14	1.58	0.00
15	1.58	0.31
16	1.58	0.53
17	1.58	0.54
18	1.58	0.11
19	1.58	0.19
20	1.58	0.30
21	1.58	0.40
22	1.58	0.75
23	1.58	0.60
24	1.58	0.12
25	1.58	0.35
26	1.58	0.76
27	1.58	0.66
28	1.58	0.25
29	1.58	1.14
30	1.58	0.55
31	1.58	0.23
32	1.58	0.54
33	1.58	0.72
34	1.58	0.70
35	1.58	1.71
36	1.58	0.43
37	1.58	0.26
38	1.58	0.38
39	1.58	1.01
40	1.58	1.28
41	1.59	2.50
42	1.59	0.18
43	1.59	1.08
44	1.59	0.78
45	1.59	1.11
46	1.59	1.67
47	1.59	1.02
48	1.59	0.42
49	1.59	0.11
50	1.59	1.93
51	1.59	0.69
52	1.59	1.32
53	1.60	2.41
54	1.60	1.05
55	1.60	0.92
56	1.60	1.47
57	1.60	2.61
58	1.61	3.40
59	1.61	1.48
60	1.61	0.40
61	1.61	1.44
62	1.61	0.55
63	1.61	1.00
64	1.61	0.67
65	1.62	1.12
66	1.62	1.96
67	1.62	0.37
68	1.62	1.18
69	1.62	1.53
70	1.62	0.45
71	1.62	0.43
72	1.62	0.57
73	1.63	1.92
74	1.63	0.83
75	1.64	2.62
76	1.65	2.27
77	1.66	1.76
78	1.67	1.49
79	1.68	1.15
80	1.68	0.67

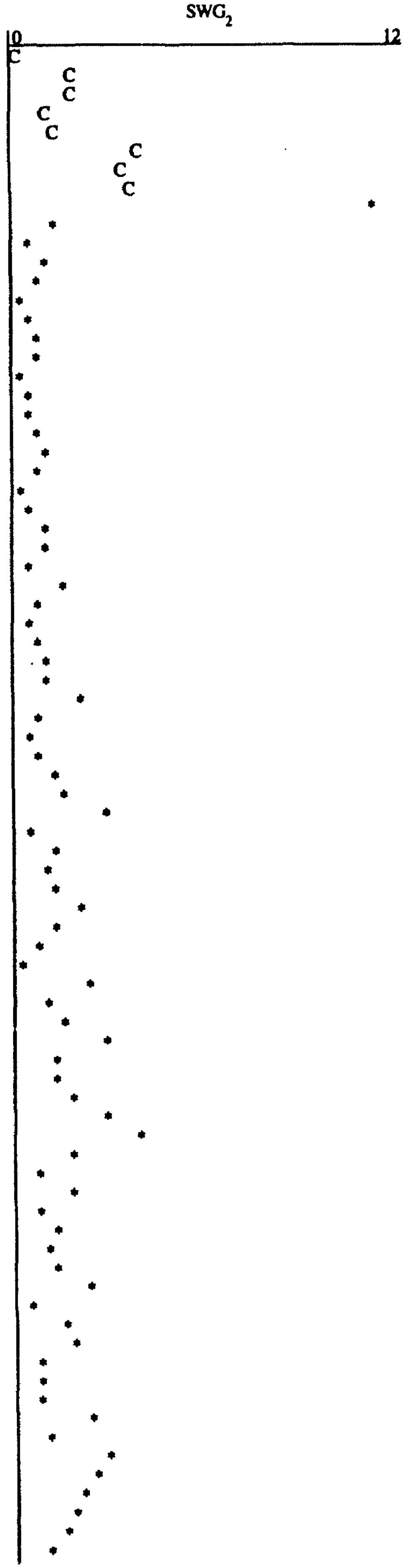


Figure 2: **Distribution of SWGs based on the ratios of the second pair of eigenvalues (REP=B R=0.5 P=10)**

	RL_2	SWG_2
1	1.00	0.00
2	1.00	0.13
3	1.00	0.13
4	1.00	0.59
5	1.00	1.00
6	1.00	0.28
7	1.00	0.21
8	1.00	1.04
9	1.07	7.92
10	1.07	1.63
11	1.07	1.60
12	1.07	0.50
13	1.07	0.81
14	1.07	0.99
15	1.07	0.32
16	1.07	1.02
17	1.07	0.45
18	1.07	0.95
19	1.07	0.33
20	1.07	0.11
21	1.07	0.24
22	1.07	1.29
23	1.07	0.76
24	1.07	0.66
25	1.07	0.10
26	1.07	0.34
27	1.07	0.79
28	1.07	1.16
29	1.07	1.00
30	1.07	0.09
31	1.07	0.07
32	1.08	0.92
33	1.08	0.70
34	1.08	0.63
35	1.08	1.12
36	1.08	0.20
37	1.08	0.50
38	1.08	0.11
39	1.08	1.37
40	1.08	1.60
41	1.08	1.20
42	1.08	1.56
43	1.08	0.65
44	1.08	1.84
45	1.08	0.82
46	1.08	0.96
47	1.08	1.79
48	1.08	0.76
49	1.08	1.22
50	1.08	1.22
51	1.08	2.00
52	1.09	0.54
53	1.09	0.90
54	1.09	1.44
55	1.09	2.48
56	1.0[illegible]	1.81
57	1.0[illegible]	0.30
58	1.09	2.32
59	1.09	1.57
60	1.09	1.06
61	1.09	1.39
62	1.10	2.02
63	1.10	1.57
64	1.10	2.41
65	1.10	0.69
66	1.10	0.52
67	1.10	0.77
68	1.10	0.26
69	1.10	0.80
70	1.10	1.03
71	1.10	0.99
72	1.10	0.34
73	1.11	1.51
74	1.11	2.19
75	1.12	2.07
76	1.13	2.98
77	1.13	1.18
78	1.13	0.29
79	1.15	1.94
80	1.17	1.62

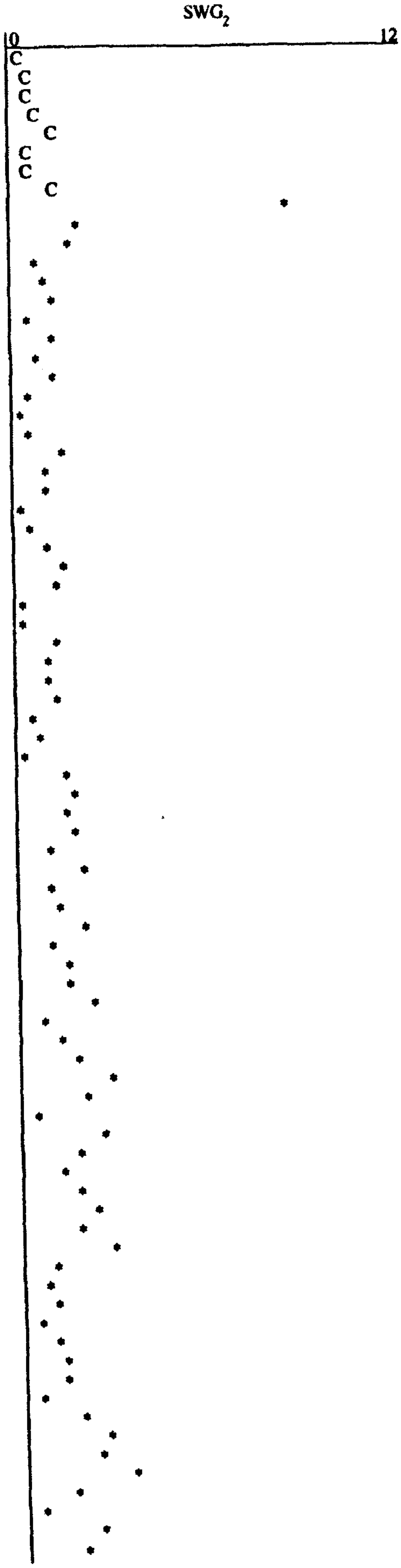


Figure 3: **Distribution of SWGs based on the ratios of the third pair of eigenvalues (REP=B R=0.7 P=10)**

	RL_3	SWG_3
1	6.10	0.00
2	6.11	0.53
3	6.21	2.57
4	6.21	0.60
5	6.41	5.15
6	6.49	3.53
7	6.49	0.52
8	6.69	6.61
9	6.83	5.93
10	7.01	7.15
11	7.04	2.90
12	7.05	1.41
13	7.06	2.23
14	7.07	1.74
15	7.07	0.93
16	7.07	0.90
17	7.08	1.28
18	7.08	1.14
19	7.08	0.80
20	7.09	1.37
21	7.09	0.94
22	7.09	0.16
23	7.09	1.10
24	7.09	0.41
25	7.09	0.39
26	7.09	1.54
27	7.09	0.59
28	7.09	0.69
29	7.10	0.70
30	7.10	0.96
31	7.10	0.84
32	7.10	0.19
33	7.10	0.58
34	7.10	1.08
35	7.10	0.85
36	7.10	0.62
37	7.10	1.11
38	7.11	1.03
39	7.11	0.63
40	7.11	0.37
41	7.11	1.56
42	7.11	0.55
43	7.11	1.57
44	7.12	1.41
45	7.12	1.78
46	7.13	1.73
47	7.13	1.65
48	7.13	0.73
49	7.13	0.22
50	7.14	2.79
51	7.15	1.64
52	7.15	2.14
53	7.15	0.50
54	7.15	0.55
55	7.16	1.46
56	7.16	0.33
57	7.16	1.51
58	7.16	1.16
59	7.17	0.92
60	7.17	0.32
61	7.17	1.83
62	7.17	0.84
63	7.17	0.38
64	7.18	1.65
65	7.19	1.64
66	7.19	0.39
67	7.20	2.22
68	7.20	1.43
69	7.22	2.05
70	7.22	1.66
71	7.24	1.81
72	7.28	3.68
73	7.30	2.11
74	7.39	4.34
75	7.50	4.60
76	7.51	1.36
77	7.53	1.82
78	7.57	1.90
79	7.77	3.69
80	8.18	3.76

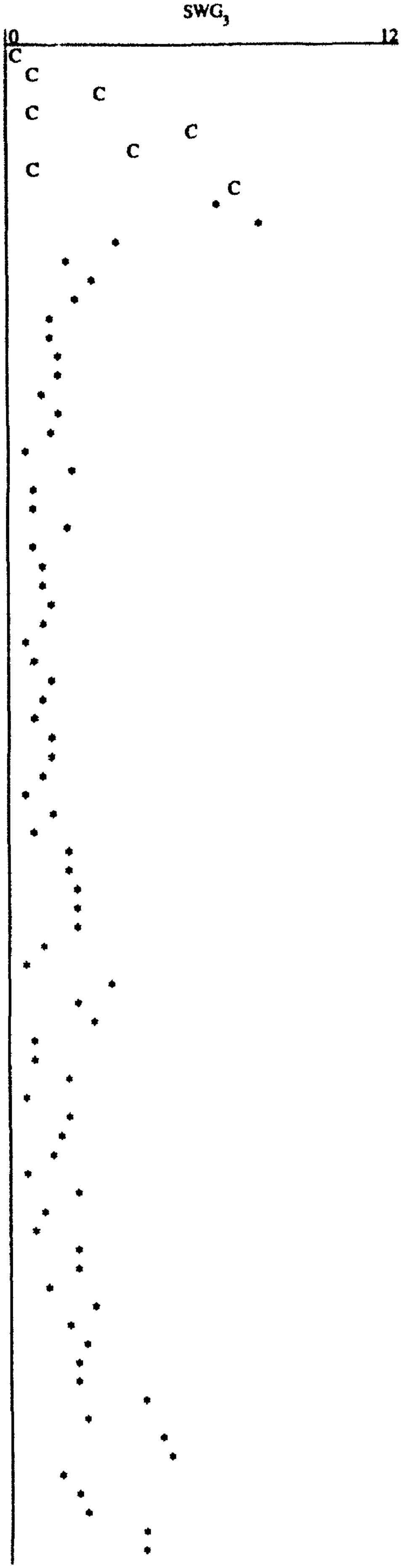


Figure 4: **Distribution of SWGs based on the ratios of the third pair of eigenvalues (REP=R R=0 P=10)**

	RL_3	SWG_3
1	8.08	0.00
2	8.11	0.94
3	8.34	3.70
4	8.75	5.84
5	9.11	6.36
6	9.21	3.81
7	9.55	7.40
8	9.72	5.78
9	10.14	9.42
10	10.17	2.43
11	10.18	1.49
12	10.18	0.36
13	10.19	1.78
14	10.19	0.45
15	10.19	0.49
16	10.19	1.07
17	10.19	0.35
18	10.19	0.56
19	10.20	0.60
20	10.20	0.92
21	10.20	0.32
22	10.20	0.33
23	10.20	0.61
24	10.20	0.59
25	10.20	0.44
26	10.20	1.04
27	10.20	0.88
28	10.20	0.65
29	10.21	1.17
30	10.21	0.51
31	10.21	0.59
32	10.21	1.17
33	10.21	0.43
34	10.21	0.51
35	10.21	0.53
36	10.21	0.43
37	10.21	0.00
38	10.21	1.20
39	10.22	1.32
40	10.22	0.94
41	10.22	1.11
42	10.22	1.49
43	10.23	1.32
44	10.23	0.27
45	10.23	0.46
46	10.23	0.55
47	10.23	1.44
48	10.24	1.86
49	10.24	0.87
50	10.24	1.58
51	10.25	1.20
52	10.25	1.53
53	10.25	1.13
54	10.25	0.69
55	10.26	1.67
56	10.27	2.57
57	10.27	0.44
58	10.29	3.01
59	10.30	2.14
60	10.30	0.19
61	10.31	2.41
62	10.33	2.79
63	10.34	1.94
64	10.37	3.23
65	10.37	1.23
66	10.46	5.69
67	10.49	3.05
68	10.50	1.68
69	10.52	2.59
70	10.53	1.62
71	10.54	1.80
72	10.56	2.06
73	10.63	3.94
74	10.66	2.01
75	10.77	4.28
76	10.87	3.68
77	10.87	0.92
78	10.92	1.97
79	10.99	2.02
80	11.97	5.34

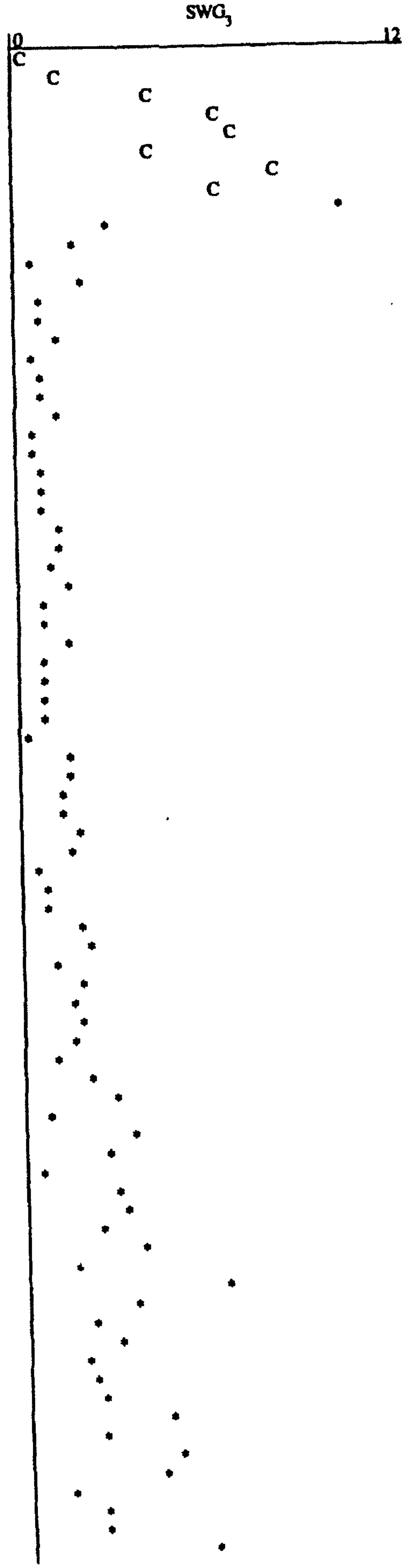


Figure 5: **Distribution of SWGs based on the ratios of the third pair of eigenvalues (REP=R R=0.5 P=10)**

	RL_3	SWG_3
1	5.37	0.00
2	5.54	2.94
3	5.55	1.10
4	5.89	7.17
5	5.91	1.85
6	6.09	6.68
7	6.28	7.38
8	6.35	4.85
9	6.63	10.39
10	6.66	3.57
11	6.67	1.06
12	6.67	1.30
13	6.68	2.08
14	6.68	0.53
15	6.68	1.02
16	6.68	0.39
17	6.69	2.00
18	6.69	0.58
19	6.69	1.23
20	6.69	0.41
21	6.69	0.46
22	6.69	0.51
23	6.69	0.30
24	6.69	0.31
25	6.69	0.43
26	6.69	0.17
27	6.69	0.57
28	6.69	0.50
29	6.69	0.86
30	6.69	1.43
31	6.70	1.52
32	6.70	1.71
33	6.70	0.75
34	6.70	0.57
35	6.70	1.58
36	6.70	0.33
37	6.70	0.38
38	6.70	0.39
39	6.70	1.12
40	6.71	1.24
41	6.71	0.79
42	6.71	1.40
43	6.71	2.22
44	6.71	0.79
45	6.72	1.21
46	6.72	1.69
47	6.72	1.37
48	6.72	0.80
49	6.72	1.76
50	6.72	0.31
51	6.73	2.25
52	6.73	0.81
53	6.73	1.29
54	6.73	1.78
55	6.74	0.95
56	6.74	0.24
57	6.74	0.93
58	6.74	1.66
59	6.74	1.63
60	6.75	1.68
61	6.78	4.74
62	6.79	3.70
63	6.80	2.64
64	6.81	1.67
65	6.83	3.64
66	6.84	2.42
67	6.85	2.96
68	6.86	2.03
69	6.88	3.28
70	6.89	2.76
71	6.93	4.01
72	6.95	2.85
73	7.05	6.07
74	7.06	1.96
75	7.07	2.26
76	7.09	1.90
77	7.10	1.19
78	7.16	2.98
79	7.39	4.85
80	7.52	2.60

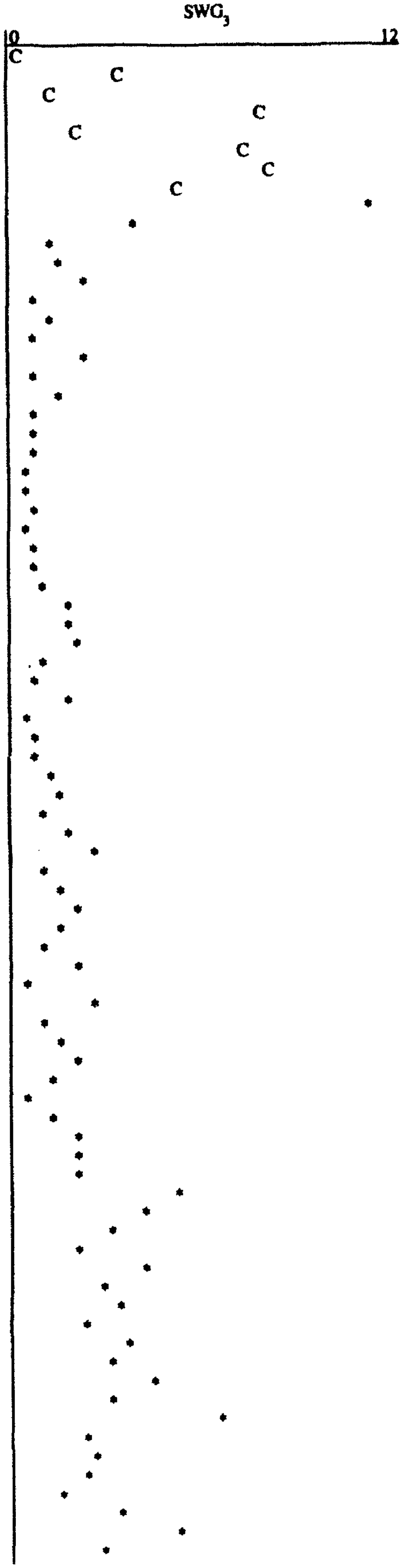


Figure 6: **Distribution of SWGs based on the ratios of the third pair of eigenvalues (REP=R R=0.7 P=10)**

	RL_3	SWG_3
1	3.32	0.00
2	3.36	1.53
3	3.40	1.94
4	3.44	2.38
5	3.65	6.42
6	3.83	6.60
7	3.83	1.02
8	3.84	1.29
9	3.95	6.43
10	3.97	2.55
11	3.97	1.36
12	3.97	0.59
13	3.97	0.44
14	3.98	2.26
15	3.98	0.90
16	3.99	1.66
17	3.99	1.17
18	3.99	1.28
19	3.99	0.48
20	3.99	0.88
21	4.00	1.35
22	4.00	1.37
23	4.00	0.53
24	4.00	0.76
25	4.00	0.77
26	4.00	0.55
27	4.00	0.00
28	4.00	0.97
29	4.00	0.80
30	4.00	1.39
31	4.00	0.81
32	4.00	0.04
33	4.01	1.16
34	4.01	0.58
35	4.01	0.58
36	4.01	0.58
37	4.01	0.12
38	4.01	0.58
39	4.01	0.84
40	4.01	0.84
41	4.01	1.02
42	4.01	0.84
43	4.01	0.16
44	4.01	1.67
45	4.01	0.81
46	4.01	1.01
47	4.02	0.83
48	4.02	1.04
49	4.02	1.52
50	4.02	1.00
51	4.02	1.52
52	4.03	1.89
53	4.03	1.13
54	4.03	1.38
55	4.03	1.25
56	4.03	1.65
57	4.04	2.44
58	4.04	1.72
59	4.05	1.50
60	4.05	1.94
61	4.05	0.96
62	4.06	2.50
63	4.06	0.86
64	4.07	2.52
65	4.08	2.09
66	4.09	2.82
67	4.10	2.21
68	4.11	2.45
69	4.12	1.95
70	4.14	3.13
71	4.14	1.68
72	4.15	0.89
73	4.20	4.56
74	4.21	1.93
75	4.22	1.18
76	4.22	0.64
77	4.25	2.35
78	4.26	1.30
79	4.35	2.94
80	4.61	3.69

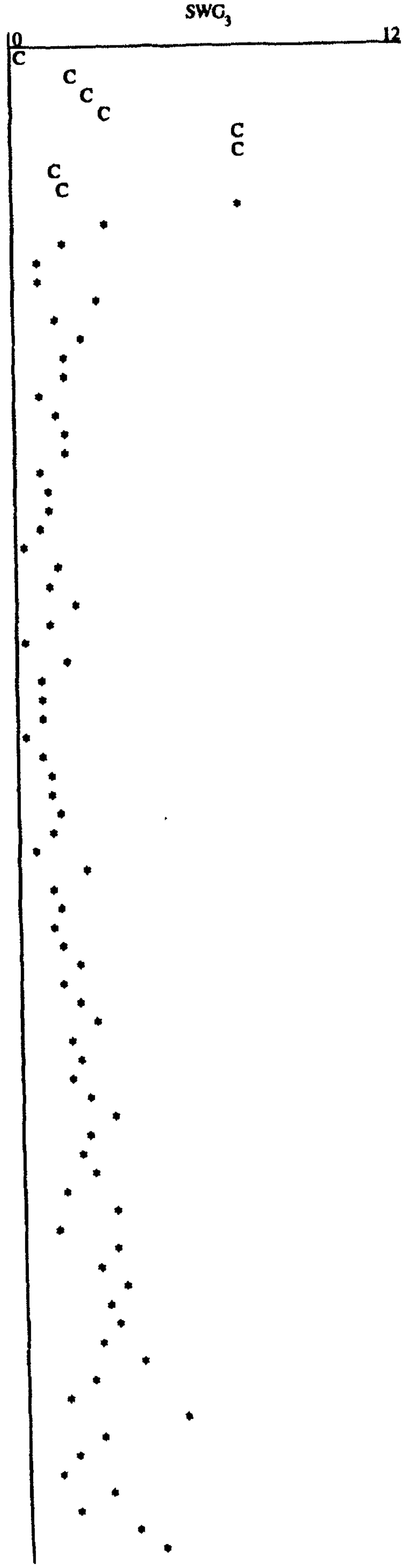


Figure 7: **Distribution of SWGs based on the ratios of the second pair of eigenvalues (REP=B R=0 P=25)**

	RL_2	SWG_2
1	3.57	0.00
2	3.58	0.57
3	3.60	1.15
4	3.61	0.59
5	3.62	1.34
6	3.65	2.37
7	3.66	1.24
8	3.67	1.80
9	3.69	1.89
10	3.70	1.88
11	3.73	3.02
12	3.76	3.36
13	3.77	1.79
14	3.77	1.51
15	3.80	3.59
16	3.81	1.30
17	3.81	1.07
18	3.82	1.99
19	3.83	2.53
20	3.89	5.57
21	3.92	4.19
22	3.92	0.28
23	3.92	0.49
24	3.92	0.29
25	3.92	0.46
26	3.92	0.81
27	3.92	0.21
28	3.92	0.48
29	3.92	0.75
30	3.92	0.52
31	3.92	0.32
32	3.92	0.87
33	3.92	0.50
34	3.93	0.83
35	3.93	0.45
36	3.93	0.89
37	3.93	0.65
38	3.93	0.83
39	3.93	0.88
40	3.93	0.38
41	3.93	0.56
42	3.93	0.94
43	3.93	1.38
44	3.93	0.64
45	3.93	0.55
46	3.94	1.96
47	3.94	1.45
48	3.95	2.61
49	3.95	1.16
50	3.96	1.77
51	3.96	0.76
52	3.96	0.37
53	3.96	0.53
54	3.96	1.43
55	3.97	2.22
56	3.98	1.91
57	3.98	0.46
58	3.98	1.77
59	3.99	1.46
60	3.99	1.30
61	4.00	2.75
62	4.01	1.87
63	4.02	2.03
64	4.02	1.25
65	4.02	1.44
66	4.03	1.73
67	4.03	0.05
68	4.04	1.63
69	4.04	0.13
70	4.04	1.48
71	4.04	0.31
72	4.05	0.89
73	4.06	1.60
74	4.08	2.38
75	4.08	0.55
76	4.13	2.94
77	4.18	2.51
78	4.19	1.22
79	4.25	2.08
80	4.28	1.07

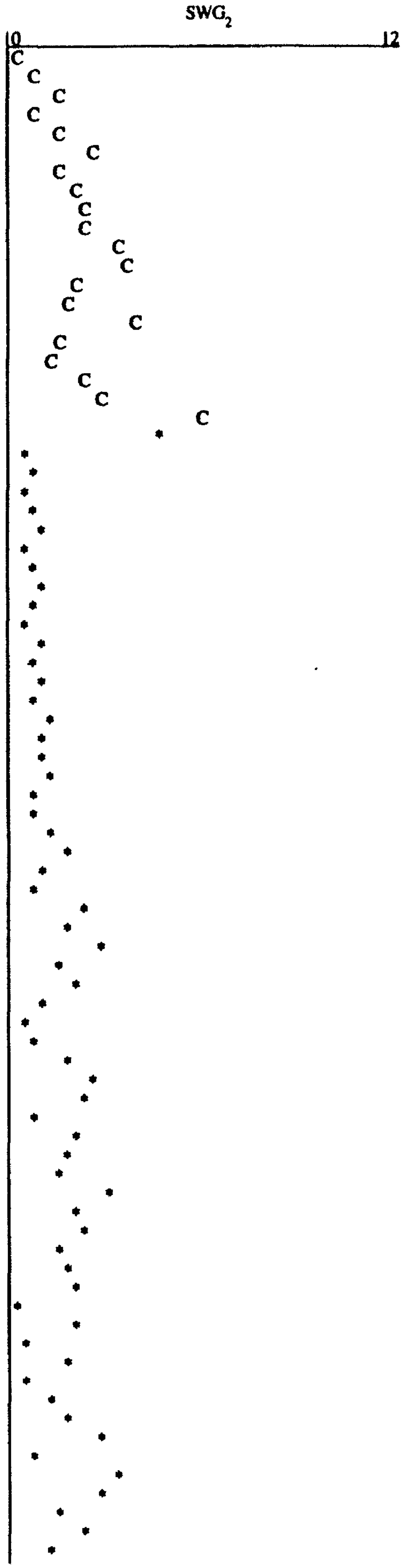


Figure 8: **Distribution of SWGs based on the ratios of the second pair of eigenvalues (REP=B R=0.5 P=25)**

	RL_2	SWG_2
1	2.42	0.00
2	2.44	1.04
3	2.45	0.89
4	2.47	1.65
5	2.49	2.18
6	2.50	1.34
7	2.51	1.87
8	2.53	2.57
9	2.53	0.91
10	2.53	0.53
11	2.53	0.37
12	2.54	1.99
13	2.54	0.66
14	2.57	3.96
15	2.58	2.85
16	2.58	1.76
17	2.58	0.88
18	2.60	3.33
19	2.62	3.98
20	2.63	2.25
21	2.63	2.44
22	2.64	2.30
23	2.64	0.71
24	2.64	0.50
25	2.64	0.79
26	2.64	0.65
27	2.64	0.52
28	2.64	0.57
29	2.64	0.38
30	2.64	0.58
31	2.64	0.37
32	2.64	1.03
33	2.65	0.97
34	2.65	0.21
35	2.65	0.42
36	2.65	0.37
37	2.65	0.64
38	2.65	0.32
39	2.65	0.22
40	2.65	0.89
41	2.65	0.52
42	2.65	0.91
43	2.66	2.96
44	2.66	0.00
45	2.66	0.51
46	2.66	2.01
47	2.66	0.32
48	2.66	0.83
49	2.66	1.92
50	2.66	0.58
51	2.67	1.13
52	2.67	1.82
53	2.67	1.26
54	2.67	0.90
55	2.67	0.56
56	2.68	2.79
57	2.68	1.97
58	2.69	1.29
59	2.69	1.85
60	2.69	0.70
61	2.69	0.82
62	2.69	1.53
63	2.70	2.09
64	2.70	1.68
65	2.70	0.29
66	2.71	1.02
67	2.71	0.52
68	2.71	0.97
69	2.72	2.14
70	2.72	1.40
71	2.72	1.17
72	2.73	1.19
73	2.73	1.07
74	2.74	2.38
75	2.77	2.69
76	2.78	1.31
77	2.80	2.38
78	2.83	2.15
79	2.84	0.97
80	2.92	2.01

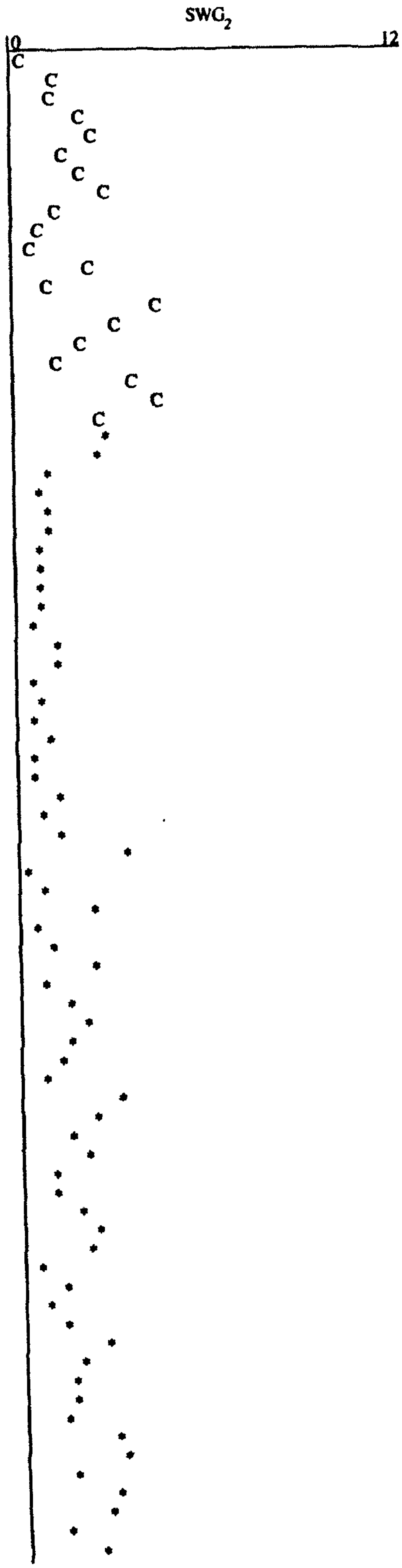


Figure 9: **Distribution of SWGs based on the ratios of the second pair of eigenvalues (REP=R R=0 P=25)**

	RL_2	SWG_2
1	2.40	0.00
2	2.40	0.31
3	2.41	0.46
4	2.41	0.57
5	2.43	1.86
6	2.44	1.29
7	2.44	0.75
8	2.46	2.31
9	2.46	0.76
10	2.46	0.80
11	2.47	1.89
12	2.48	1.59
13	2.49	1.94
14	2.49	0.90
15	2.49	0.41
16	2.49	1.11
17	2.51	3.60
18	2.52	1.57
19	2.53	2.26
20	2.56	4.44
21	2.60	5.17
22	2.60	0.96
23	2.60	0.81
24	2.60	0.49
25	2.60	0.57
26	2.60	0.84
27	2.60	0.30
28	2.61	0.11
29	2.61	0.28
30	2.61	0.99
31	2.61	0.76
32	2.61	0.63
33	2.61	0.72
34	2.61	0.33
35	2.61	0.55
36	2.61	0.57
37	2.61	0.78
38	2.61	0.35
39	2.61	0.84
40	2.61	0.28
41	2.61	0.13
42	2.61	0.97
43	2.62	1.69
44	2.62	0.95
45	2.62	0.96
46	2.62	1.38
47	2.62	1.36
48	2.62	0.36
49	2.62	1.34
50	2.63	1.59
51	2.63	1.09
52	2.63	2.13
53	2.65	3.33
54	2.65	1.03
55	2.65	0.04
56	2.65	1.86
57	2.66	0.54
58	2.66	1.07
59	2.66	0.84
60	2.66	0.54
61	2.66	0.38
62	2.66	1.24
63	2.66	0.09
64	2.66	1.18
65	2.66	0.56
66	2.66	0.19
67	2.66	0.51
68	2.67	0.81
69	2.68	2.32
70	2.69	2.55
71	2.69	0.45
72	2.70	0.86
73	2.70	1.03
74	2.70	0.85
75	2.71	1.35
76	2.73	1.95
77	2.73	0.86
78	2.76	1.69
79	2.77	1.09
80	2.81	1.21

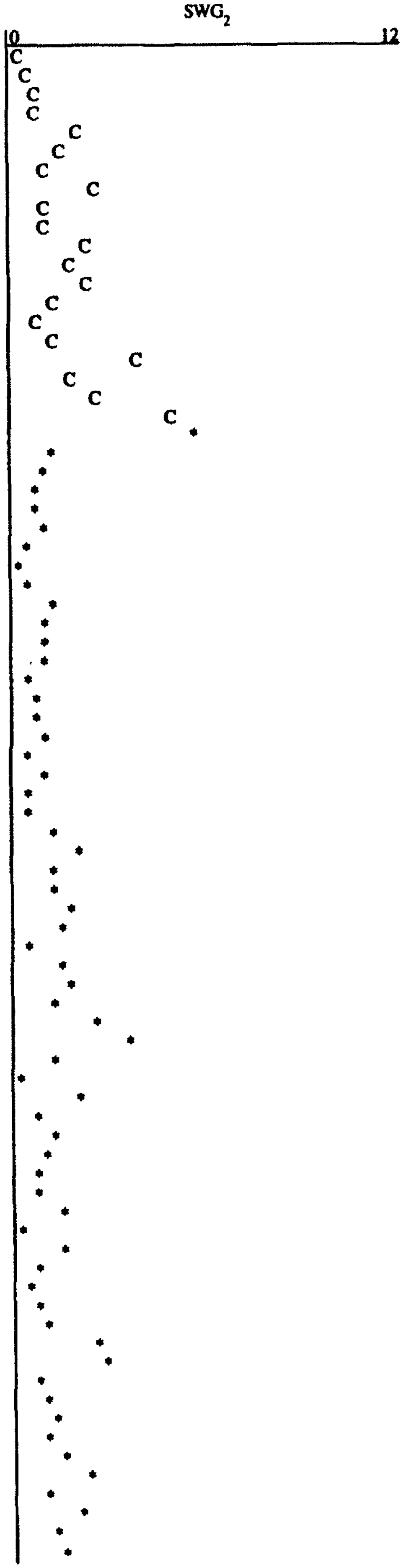


Figure 10: Distribution of SWGs based on the ratios of the second pair of eigenvalues (REP=R R=0.5 P=25)

	RL_2	SWG_2
1	1.61	0.00
2	1.63	1.24
3	1.64	1.61
4	1.65	0.88
5	1.65	1.25
6	1.65	1.03
7	1.65	0.88
8	1.66	0.97
9	1.66	0.66
10	1.66	1.60
11	1.66	1.03
12	1.67	3.29
13	1.67	0.98
14	1.67	0.88
15	1.68	2.15
16	1.68	1.87
17	1.68	1.62
18	1.69	3.04
19	1.69	2.22
20	1.73	6.79
21	1.73	1.20
22	1.73	1.86
23	1.73	0.31
24	1.73	1.64
25	1.73	0.36
26	1.73	0.76
27	1.73	1.27
28	1.73	1.14
29	1.73	0.28
30	1.73	0.23
31	1.74	1.35
32	1.74	0.29
33	1.74	1.57
34	1.74	0.77
35	1.74	1.45
36	1.74	0.79
37	1.74	1.10
38	1.74	0.60
39	1.74	1.74
40	1.74	1.67
41	1.74	1.08
42	1.74	0.31
43	1.74	1.05
44	1.74	1.28
45	1.74	0.52
46	1.75	1.66
47	1.75	0.79
48	1.75	0.46
49	1.75	1.32
50	1.75	1.26
51	1.75	0.08
52	1.75	1.41
53	1.75	1.13
54	1.75	0.67
55	1.75	0.85
56	1.75	0.46
57	1.75	1.48
58	1.75	1.00
59	1.76	1.82
60	1.76	1.13
61	1.76	3.59
62	1.77	0.87
63	1.77	2.65
64	1.77	0.34
65	1.77	1.17
66	1.77	1.03
67	1.77	1.26
68	1.78	1.89
69	1.78	1.49
70	1.78	1.53
71	1.79	2.56
72	1.81	3.68
73	1.81	0.80
74	1.81	0.29
75	1.81	1.08
76	1.82	2.21
77	1.83	1.26
78	1.83	1.02
79	1.83	0.45
80	1.86	1.75

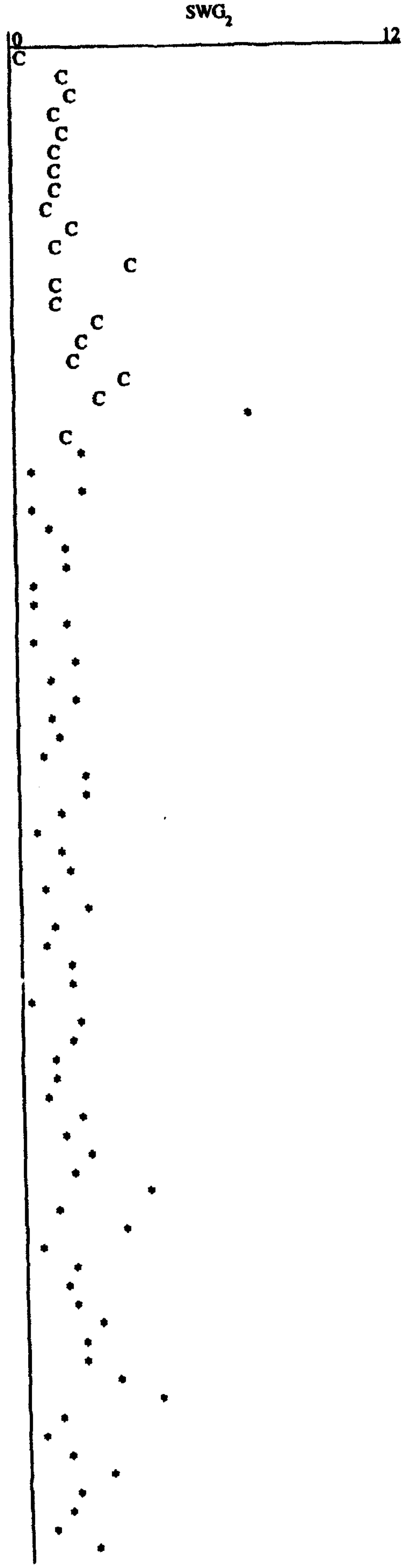


Figure 11: **ROC curves for the ratio of the first and second pairs of eigenvalues REP=B R=0.0 P=10**

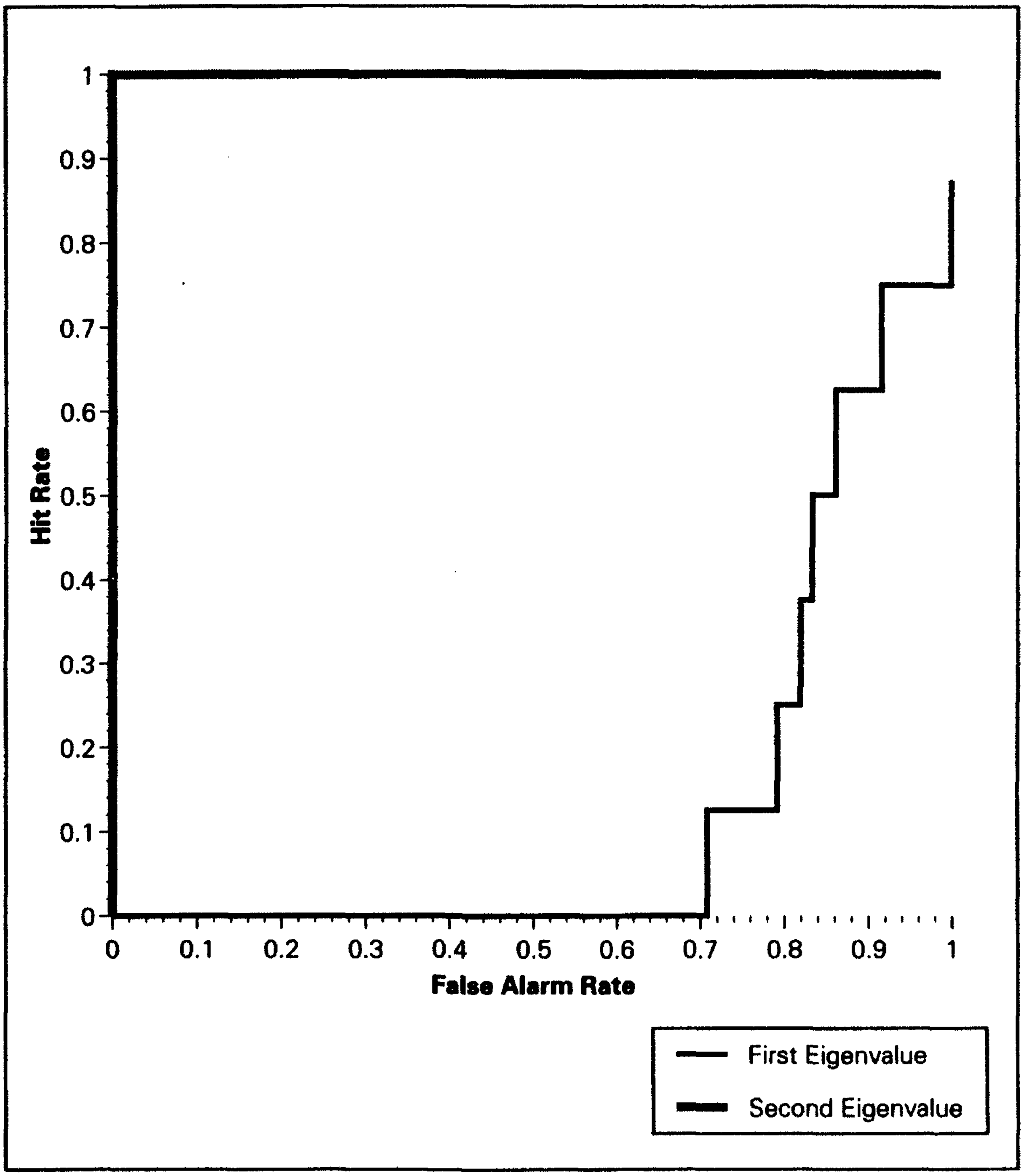


Figure 12: **ROC curves for the ratio of the first and second pairs of eigenvalues REP=B R=0.5 P=10**

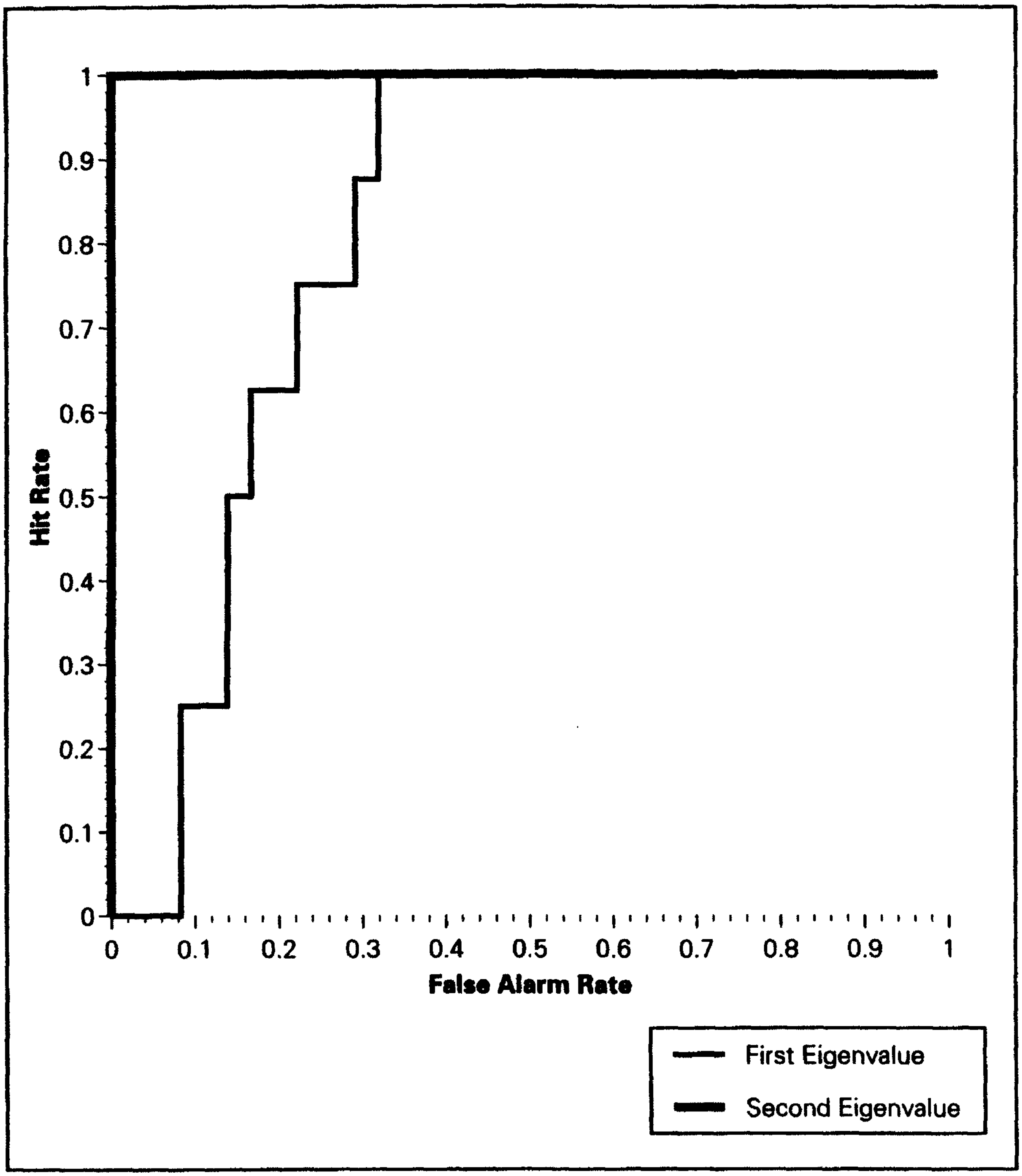


Figure 13: **ROC curves for the ratio of the first and third pairs of eigenvalues REP=B R=0.7 P=10**

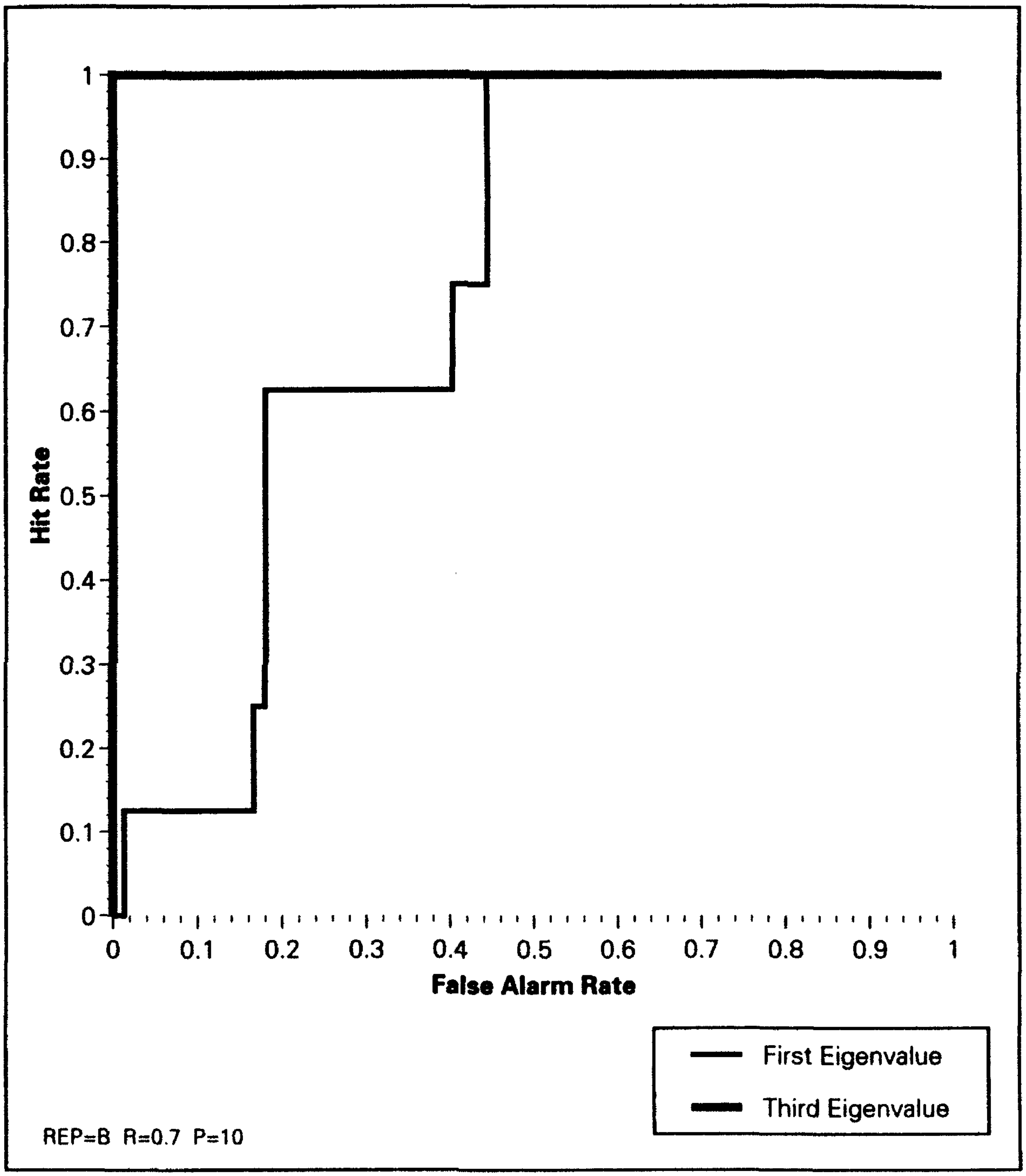


Figure 14: **ROC curves for the ratio of the first and third pairs of eigenvalues REP=R R=0.0 P=10**

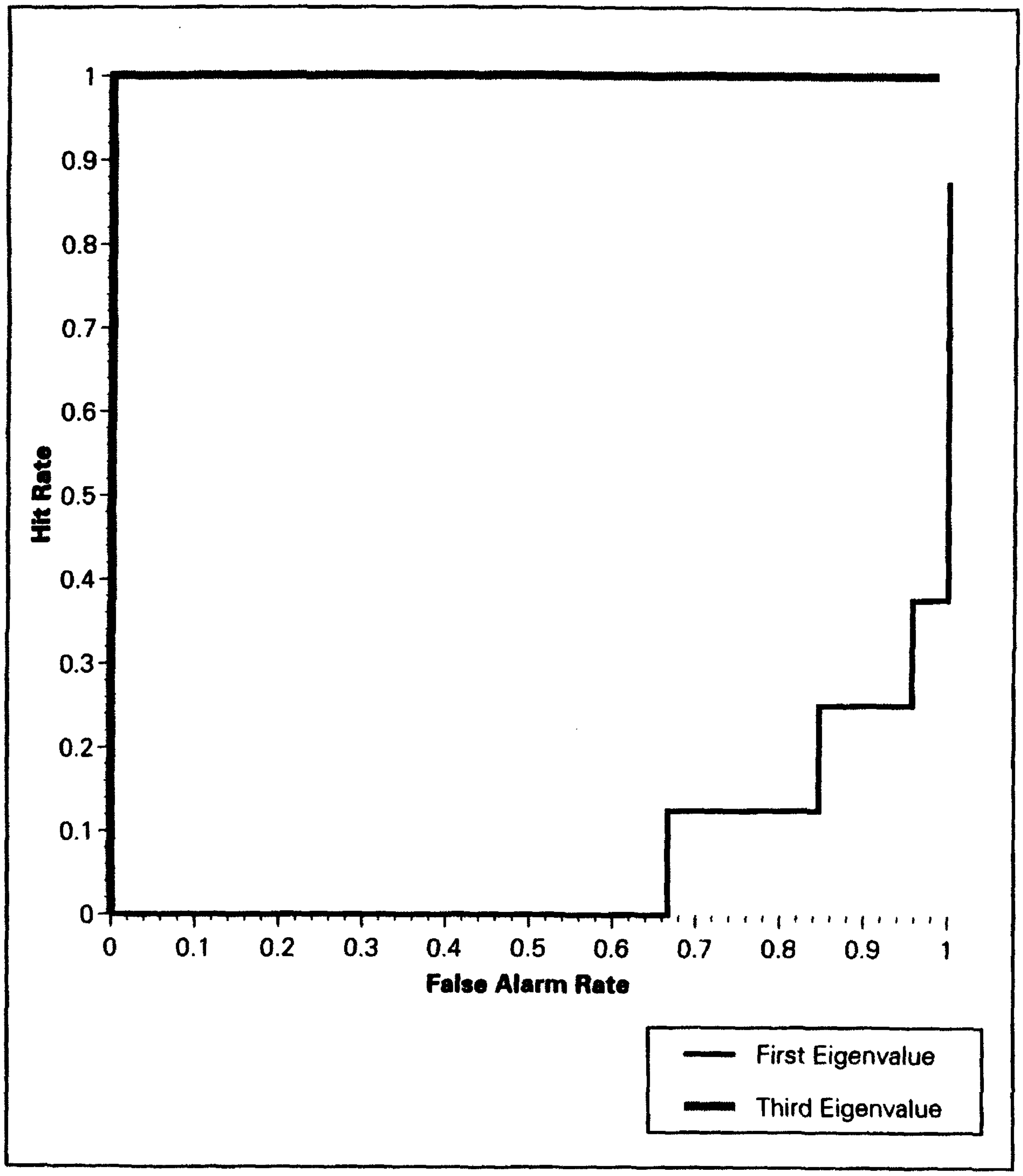


Figure 15: **ROC curves for the ratio of the first and third pairs of eigenvalues REP=R R=0.5 P=10**

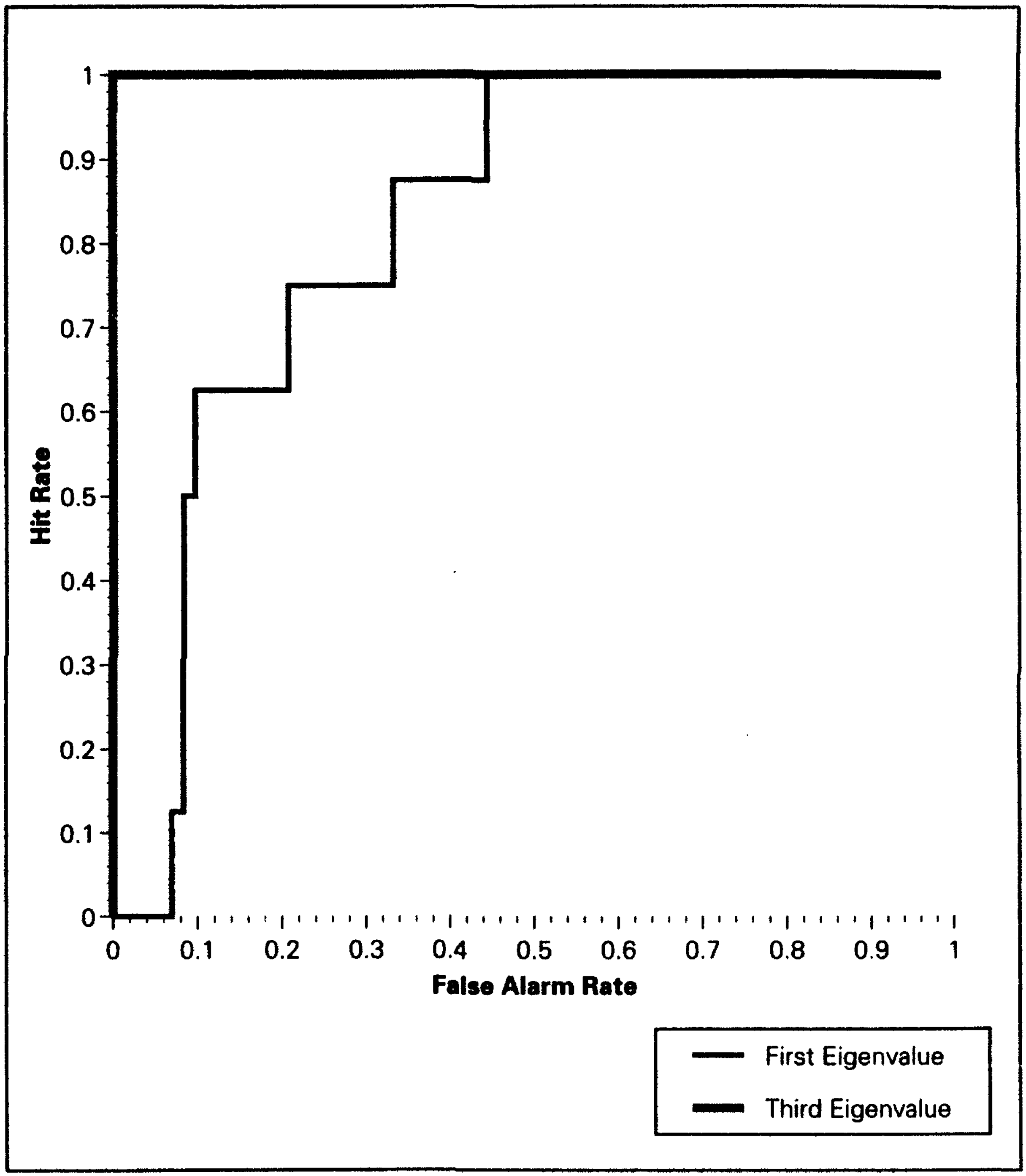


Figure 16: **ROC curves for the ratio of the first and third pairs of eigenvalues REP=R R=0.7 P=10**

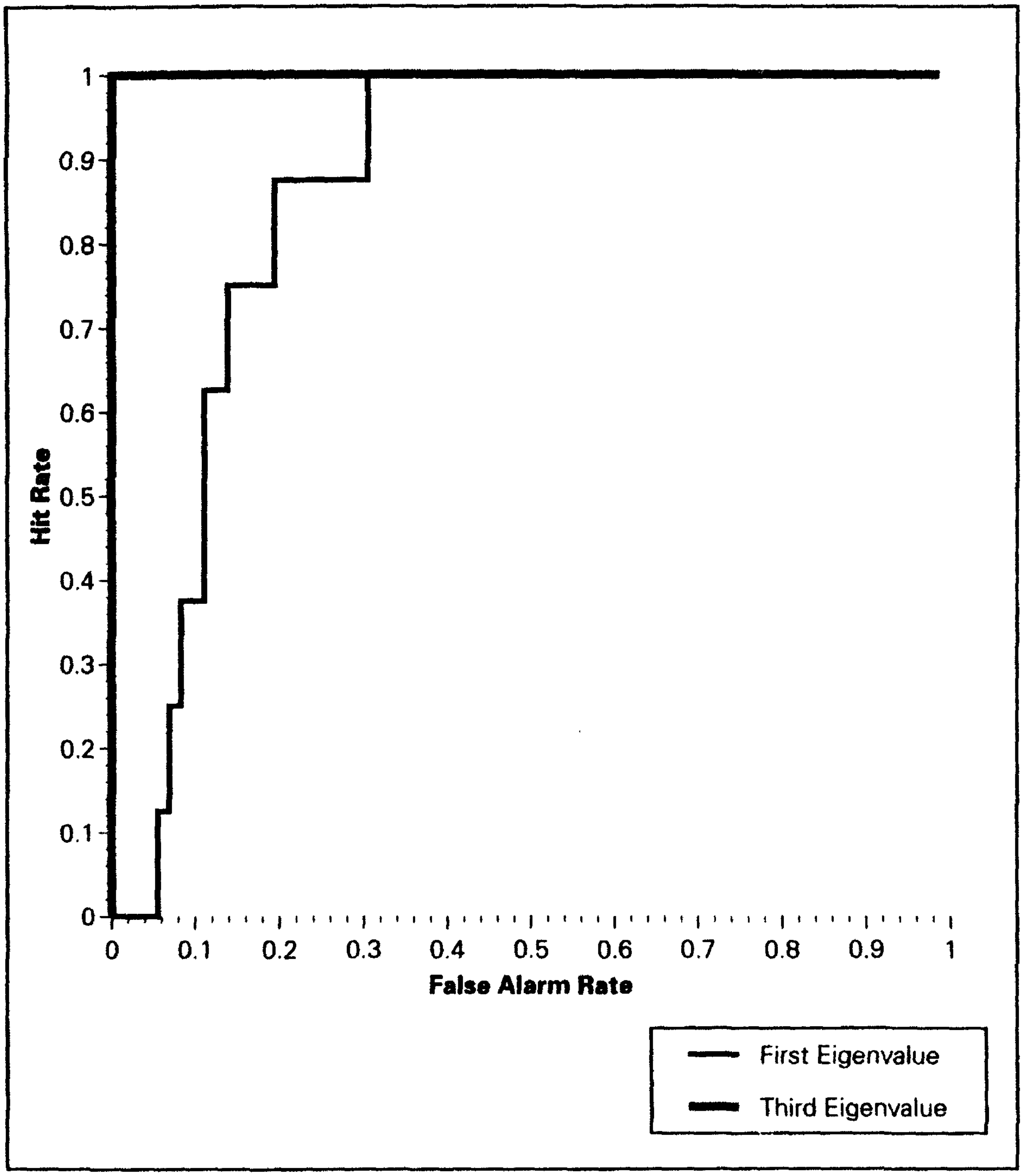


Figure 17: **ROC curves for the ratio of the first and second pairs of eigenvalues REP=B R=0.0 P=25**

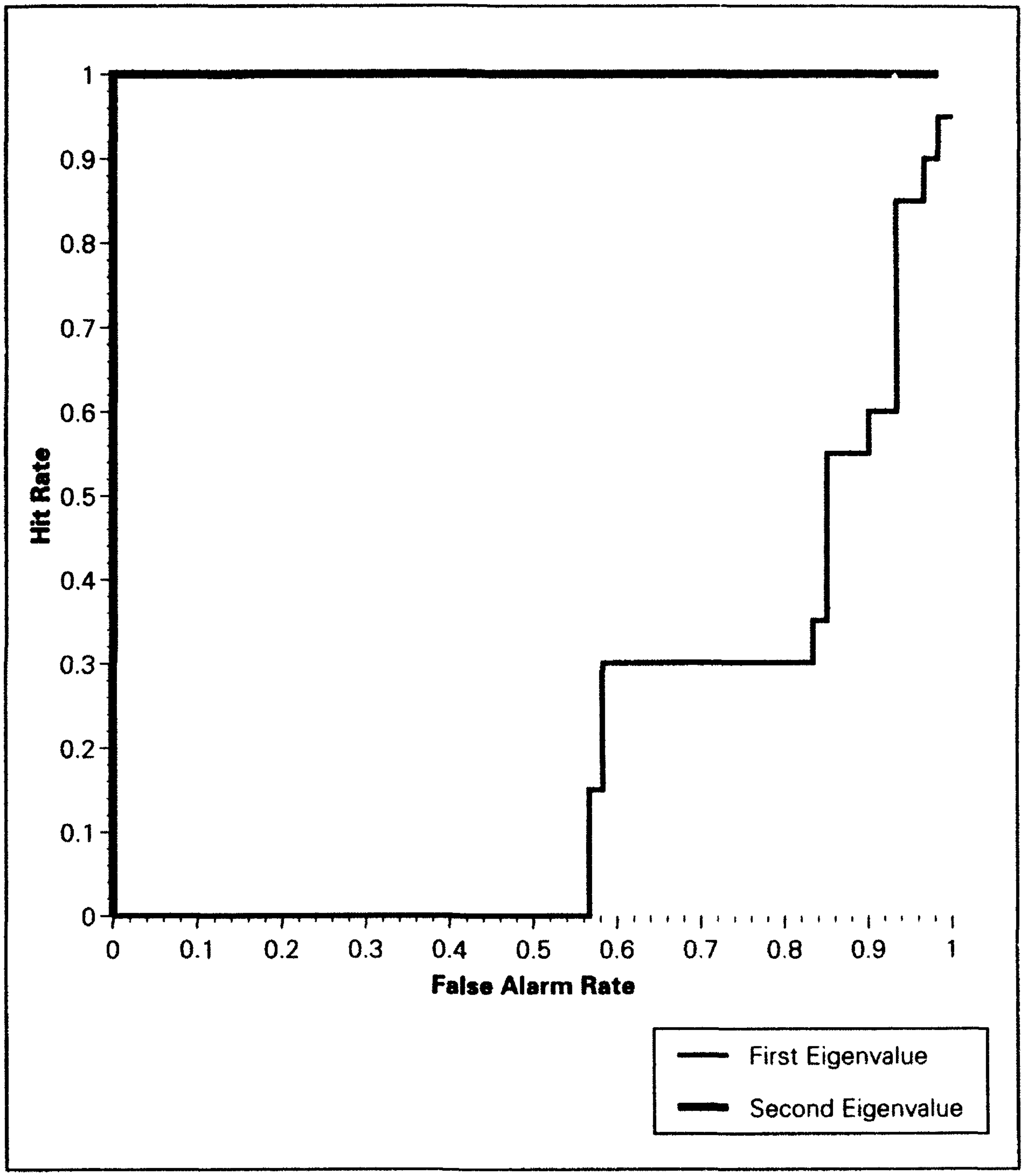


Figure 18: **ROC curves for the ratio of the first and second pairs of eigenvalues REP=B R=0.5 P=25**

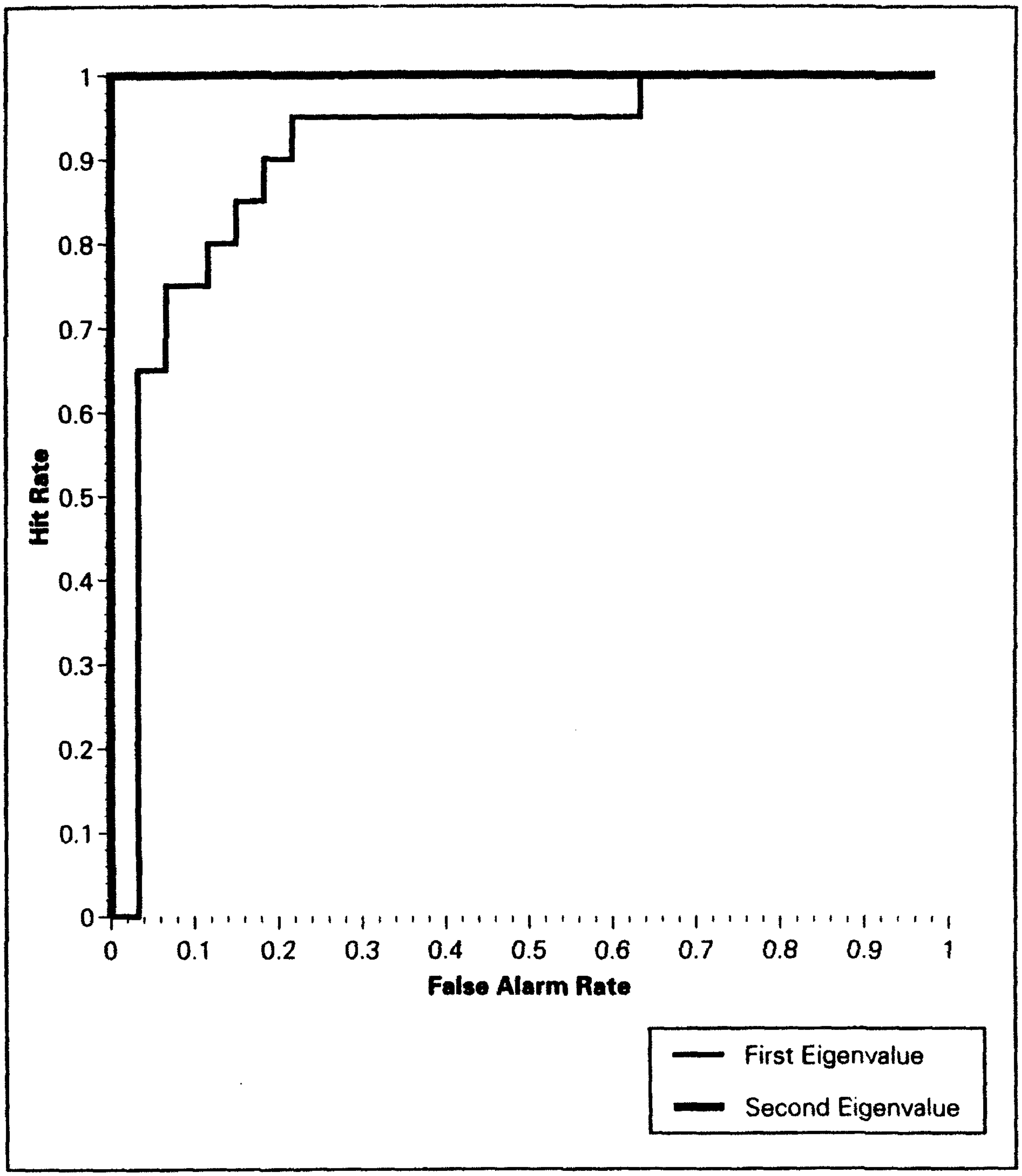


Figure 19: **ROC curves for the ratio of the first and second pairs of eigenvalues REP=R R=0.0 P=25**

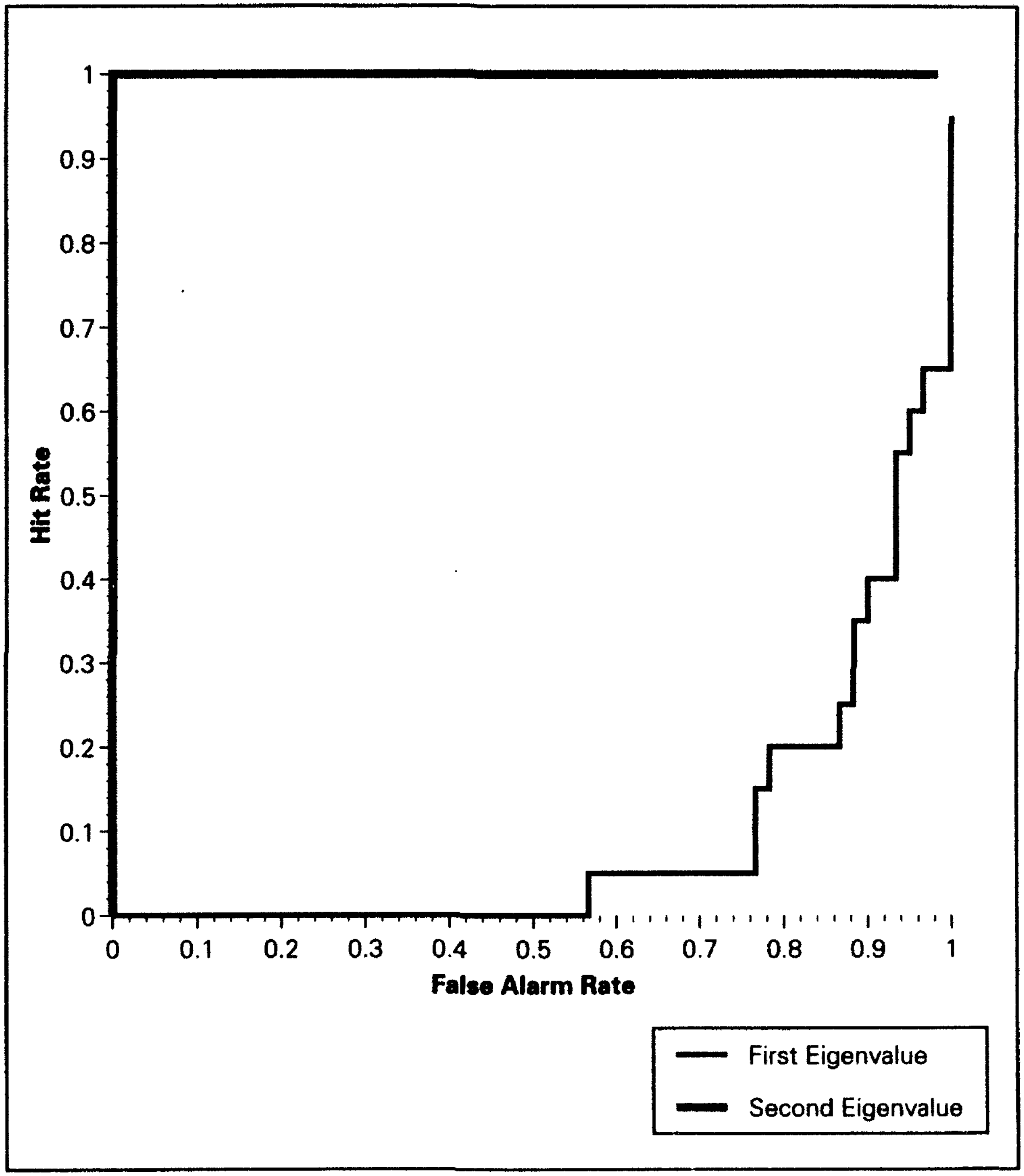
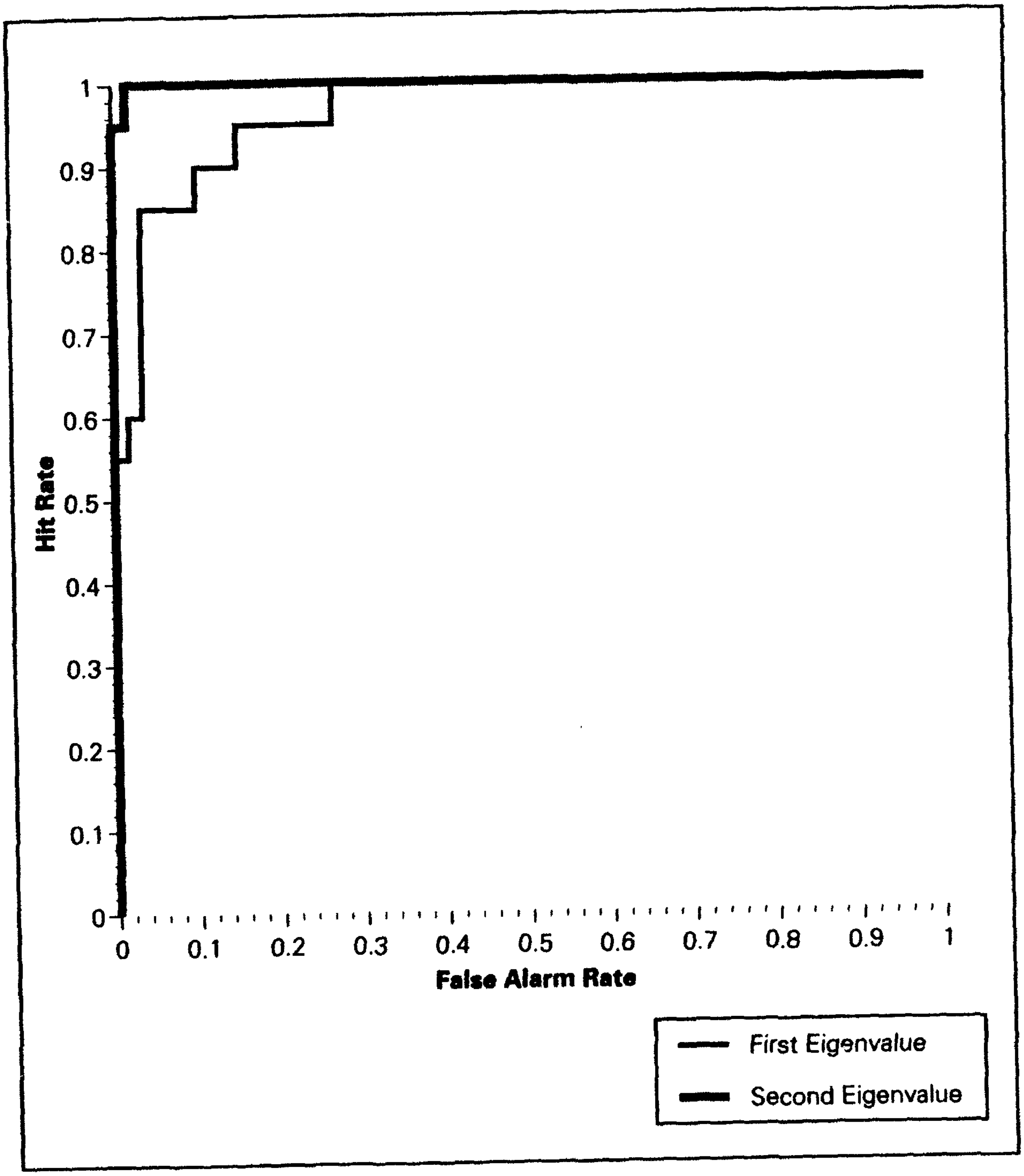


Figure 20: **ROC curves for the ratio of the first and second pairs of eigenvalues REP=R R=0.5 P=25**



APPENDIX 1:

Parameters of the items used in the R tests

Item	a	b	c
001	0.70	-1.97	0.00
002	1.01	-1.26	0.08
003	1.19	-1.15	0.02
004	0.81	-0.64	0.28
005	0.57	-1.41	0.03
006	0.91	-1.07	0.00
007	0.60	-1.50	0.06
008	0.54	-1.08	0.20
009	1.25	-2.06	0.03
010	1.30	-1.45	0.00
011	1.13	-1.21	0.02
012	1.21	-0.89	0.09
013	1.01	-0.67	0.11
014	1.82	-0.54	0.16
015	1.36	-0.67	0.14
016	1.57	-0.06	0.34
017	2.02	-0.79	0.23
018	0.72	-0.74	0.16
019	1.06	-0.63	0.09
020	1.50	0.08	0.22
021	1.57	0.93	0.19
022	1.23	1.19	0.19
023	0.86	0.53	0.27
024	0.81	0.65	0.27
025	1.38	0.11	0.20
026	0.72	0.22	0.27
027	1.70	0.34	0.22
028	1.55	0.48	0.27
029	1.35	0.46	0.30
030	1.79	0.98	0.19
031	1.62	0.41	0.17
032	2.44	0.84	0.19
033	2.44	1.02	0.17
034	2.44	1.61	0.31
035	2.44	0.87	0.23
036	1.40	0.53	0.20
037	0.81	0.41	0.20
038	1.43	0.79	0.20
039	1.23	1.34	0.23
040	0.79	1.39	0.17
041	0.57	-1.87	0.03
042	0.80	-1.66	0.02
043	1.07	-1.38	0.00
044	0.85	-1.06	0.00
045	1.32	-0.17	0.28
046	1.10	-1.07	0.08
047	0.71	-0.67	0.08
048	1.31	-1.17	0.02
049	1.73	-1.22	0.03
050	1.60	-1.29	0.03
051	1.03	-1.13	0.03
052	1.21	-0.90	0.00
053	0.63	-0.02	0.09
054	1.14	-0.19	0.19
055	0.85	-0.83	0.08
056	1.14	0.20	0.16
057	0.71	0.86	0.19
058	1.36	-0.81	0.06
059	0.95	-0.27	0.09
060	1.29	0.43	0.31
061	0.85	0.28	0.13
062	1.89	0.88	0.17
063	0.73	0.77	0.17
064	0.97	0.56	0.30
065	1.41	0.45	0.16
066	1.82	0.66	0.30
067	1.70	-0.06	0.19
068	1.70	0.35	0.28
069	1.48	0.99	0.16
070	1.72	0.93	0.19
071	1.80	0.75	0.16
072	2.47	1.09	0.22
073	1.56	1.12	0.31
074	2.04	1.20	0.33
075	1.99	1.29	0.22
076	1.09	0.80	0.27
077	1.10	0.61	0.17
078	1.82	1.19	0.22
079	2.47	1.24	0.30
080	2.04	1.61	0.17

APPENDIX 2:

Parameters of the items used in the B tests

Item	a	b	c
1	1.41	0.42	0.25
2	1.00	0.06	0.28
3	1.19	0.40	0.20
4	1.35	1.61	0.27
5	1.35	-0.30	0.25
6	1.53	1.15	0.14
7	1.45	0.60	0.10
8	1.09	0.68	0.22
9	0.93	0.13	0.28
10	0.89	-1.01	0.20
11	1.25	-0.51	0.23
12	1.37	-0.90	0.30
13	0.88	-0.40	0.22
14	1.30	-0.21	0.20
15	0.81	0.77	0.22
16	1.02	-0.69	0.28
17	0.96	0.81	0.11
18	0.99	0.48	0.28
19	1.42	0.84	0.12
20	0.62	0.36	0.20
21	1.19	1.17	0.21
22	1.14	0.66	0.10
23	0.96	-0.02	0.18
24	1.23	-0.19	0.19
25	0.89	-0.20	0.13
26	1.12	-0.79	0.20
27	0.65	0.72	0.26
28	0.83	-0.29	0.17
29	0.90	-1.14	0.24
30	1.27	-0.85	0.28
31	0.90	-0.25	0.24
32	1.32	-1.76	0.24
33	1.42	-0.28	0.18
34	1.45	1.44	0.10
35	1.30	1.48	0.14
36	1.17	-0.11	0.29
37	1.39	1.99	0.20
38	0.78	1.73	0.14
39	1.48	-0.10	0.23
40	1.31	0.72	0.12
41	1.54	0.10	0.15
42	1.54	1.50	0.17
43	1.13	0.10	0.24
44	1.27	-0.39	0.10
45	0.84	-1.11	0.29
46	1.21	-0.12	0.25
47	1.06	0.53	0.28
48	1.19	-1.32	0.22
49	0.85	-0.42	0.14
50	0.65	0.81	0.23
51	1.32	1.30	0.27
52	1.30	1.20	0.14
53	0.79	-1.49	0.24
54	1.07	-1.96	0.17
55	1.31	0.47	0.10
56	0.93	1.25	0.19
57	1.08	0.69	0.28
58	1.27	-0.18	0.14
59	1.26	1.02	0.20
60	1.06	0.41	0.20
61	0.93	0.10	0.27
62	1.24	-0.10	0.16
63	1.20	-0.21	0.12
64	1.40	0.19	0.14
65	0.93	0.92	0.12
66	1.08	1.60	0.17
67	0.75	0.30	0.14
68	1.26	-0.13	0.19
69	1.04	-0.26	0.27
70	0.74	1.38	0.20
71	1.44	-1.51	0.26
72	0.89	0.38	0.18
73	0.73	0.85	0.28
74	0.95	0.25	0.24
75	0.77	-0.16	0.22
76	0.86	-1.70	0.22
77	1.29	0.74	0.29
78	1.55	0.14	0.19
79	1.35	-0.39	0.20
80	1.26	0.80	0.17

APPENDIX 3:

Mean values (and standard deviations) of the estimates of parameter a for items loaded on the dominant (main) and the nuisance (cont.) dimension

	p = 10		p = 25		p = 50	
	main (n=72)	cont. (n=8)	main (n=60)	cont. (n=20)	main (n=40)	cont. (n=40)
Rep = R						
True	1.31 (.51)	1.45 (.56)	1.31 (.50)	1.38 (.55)	1.29 (.48)	1.36 (.54)
p=0	1.24 (.47)	1.37 (.47)	1.23 (.45)	1.30 (.55)	1.20 (.42)	1.29 (.52)
r=.7	1.24 (.50)	.70 (.15	1.20 (.45)	.80 (.22)	.97 (.29)	1.06 (.38)
r=.5	1.23 (.48)	.53 (.12)	1.21 (.45)	.60 (.14)	.86 (.23)	.96 (.34)
r=.0	1.25 (.49)	.57 (.77)	1.25 (.51)	.74 (.88)	.27 (.01)	1.17 (.47)

Rep = B						
True	1.13 (.25)	1.05 (.21)	1.12 (.25)	1.12 (.24)	1.13 (.25)	1.12 (.24)
p=0	1.03 (.26)	.92 (.16)	1.02 (.26)	1.02 (.25)	1.05 (.28)	.99 (.23)
r=.7	1.03 (.25)	.57 (.11)	.99 (.23)	.58 (.11)	.82 (.21)	.83 (.17)
r=.5	1.03 (.25)	.44 (.06)	1.00 (.25)	.51 (.10)	.83 (.16)	.67 (.14)
r=.0	1.02 (.28)	.28 (.0)	1.00 (.26)	.93 (.99)	.92 (.19)	.27 (.02)

APPENDIX 4:

Mean values (and standard deviations) of item reliabilities for items loaded on the dominant (main) and the nuisance (cont.) dimension

	p = 10		p = 25		p = 50	
	main (n=72)	cont. (n=8)	main (n=60)	cont. (n=20)	main (n=40)	cont. (n=40)
Rep = R						
True	.37 (.13)	.44 (.15)	.37 (.14)	.36 (.12)	.39 (.13)	.35 (.14)
p=0	.34 (.14)	.42 (.16)	.36 (.15)	.35 (.13)	.37 (.14)	.34 (.15)
r=.7	.36 (.14)	.22 (.11)	.36 (.15)	.21 (.08)	.33 (.13)	.28 (.11)
r=.5	.36 (.14)	.12 (.06)	.36 (.15)	.12 (.04)	.29 (.12)	.24 (.10)
r=.0	.37 (.14)	.03 (.02)	.36 (.15)	.02 (.01)	.04 (.01)	.30 (.13)
Rep = B						
True	.34 (.12)	.37 (.06)	.35 (.11)	.31 (.13)	.35 (.12)	.33 (.11)
p=0	.32 (.12)	.35 (.05)	.33 (.11)	.30 (.14)	.34 (.12)	.31 (.12)
r=.7	.33 (.12)	.18 (.05)	.32 (.12)	.17 (.09)	.27 (.11)	.26 (.10)
r=.5	.33 (.12)	.10 (.01)	.32 (.12)	.11 (.04)	.22 (.10)	.25 (.08)
r=.0	.33 (.12)	.02 (.01)	.32 (.12)	.01 (.01)	.04 (.02)	.28 (.11)

APPENDIX 5a:

Revised Modified Parallel Analysis (RMPA) of 18 tests: Proportion of Standardized Weighted Gaps (SWGs) above the 2.25 detection threshold

First EV.	P 10	25	50	Mean
R				
0	0.025	0.025	0.019	0.023
5	0.000	0.013	0.013	0.009
7	0.025	0.006	0.000	0.010
Mean	0.017	0.015	0.011	0.014

Second EV.	P 10	25	50	Mean
R				
0	0.101	0.127	0.158	0.129
5	0.057	0.120	0.051	0.076
7	0.063	0.082	0.038	0.061
Mean	0.074	0.110	0.082	0.089

Third EV.	P 10	25	50	Mean
R				
0	0.209	0.146	0.241	0.199
5	0.171	0.133	0.209	0.171
7	0.184	0.139	0.152	0.158
Mean	0.188	0.139	0.201	0.176

APPENDIX 5b:

Revised Modified Parallel Analysis (RMPA) of 18 tests: Proportion of Standardized Weighted Gaps (SWGs) above 95th empirical percentile

First EV.	P			
	10	25	50	Mean
R				
0	0.051	0.051	0.051	0.051
5	0.051	0.051	0.051	0.051
7	0.051	0.051	0.051	0.051
Mean	0.051	0.051	0.051	0.051

Second EV.	P			
	10	25	50	Mean
R				
0	0.146	0.241	0.184	0.190
5	0.139	0.184	0.114	0.146
7	0.089	0.133	0.082	0.101
Mean	0.125	0.186	0.127	0.146

Third EV.	P			
	10	25	50	Mean
R				
0	0.335	0.222	0.184	0.247
5	0.348	0.158	0.228	0.245
7	0.228	0.184	0.234	0.215
Mean	0.304	0.188	0.215	0.236

APPENDIX 5c:

Revised Modified Parallel Analysis (RMPA) of 18 tests: Proportion of Standardized Weighted Gaps (SWGs) above 95th percentile (Chebyshev inequality + unimodality)

First EV.	P			
	10	25	50	Mean
R				
0	0.006	0.019	0.019	0.015
5	0.000	0.013	0.006	0.006
7	0.006	0.006	0.000	0.004
Mean	0.004	0.013	0.008	0.008

Second EV.	P			
	10	25	50	Mean
R				
0	0.057	0.076	0.152	0.095
5	0.057	0.095	0.032	0.061
7	0.025	0.044	0.019	0.029
Mean	0.046	0.072	0.068	0.062

Third EV.	P			
	10	25	50	Mean
R				
0	0.171	0.089	0.209	0.156
5	0.184	0.095	0.114	0.131
7	0.127	0.114	0.120	0.120
Mean	0.161	0.099	0.148	0.136

APPENDIX 5d:

Revised Modified Parallel Analysis (RMPA) of 18 tests: Proportion of Standardized Weighted Gaps (SWGs) above 95th percentile (Chebyshev inequality)

First EV.	P			
	10	25	50	Mean
R				
0	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000
Mean	0.000	0.000	0.000	0.000

Second EV.	P			
	10	25	50	Mean
R				
0	0.006	0.044	0.101	0.050
5	0.013	0.032	0.000	0.015
7	0.006	0.006	0.000	0.004
Mean	0.008	0.027	0.034	0.023

Third EV.	P			
	10	25	50	Mean
R				
0	0.076	0.032	0.089	0.066
5	0.095	0.044	0.057	0.065
7	0.089	0.057	0.013	0.053
Mean	0.087	0.044	0.053	0.061

APPENDIX 5e:

Revised Modified Parallel Analysis (RMPA) of 18 tests: Proportion of Standardized Weighted Gaps (SWGs) above 99th empirical percentile

First EV.	P			
	10	25	50	Mean
R				
0	0.013	0.013	0.013	0.013
5	0.013	0.013	0.013	0.013
7	0.013	0.013	0.013	0.013
Mean	0.013	0.013	0.013	0.013

Second EV.	P			
	10	25	50	Mean
R				
0	0.044	0.063	0.127	0.078
5	0.108	0.070	0.044	0.074
7	0.013	0.044	0.057	0.038
Mean	0.055	0.059	0.076	0.063

Third EV.	P			
	10	25	50	Mean
R				
0	0.146	0.076	0.139	0.120
5	0.272	0.108	0.127	0.169
7	0.114	0.114	0.171	0.133
Mean	0.177	0.099	0.146	0.141

APPENDIX 5f:

Revised Modified Parallel Analysis (RMPA) of 18 tests: Proportion of Standardized Weighted Gaps (SWGs) above 99th percentile (Chebyshev inequality + unimodality)

First EV.	P			
	10	25	50	Mean
R				
0	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000
Mean	0.000	0.000	0.000	0.000

Second EV.	P			
	10	25	50	Mean
R				
0	0.006	0.013	0.044	0.021
5	0.006	0.006	0.000	0.004
7	0.000	0.000	0.000	0.000
Mean	0.004	0.006	0.015	0.008

Third EV.	P			
	10	25	50	Mean
R				
0	0.044	0.000	0.057	0.034
5	0.051	0.013	0.006	0.023
7	0.057	0.006	0.006	0.023
Mean	0.051	0.006	0.023	0.027

APPENDIX 5g:

Revised Modified Parallel Analysis (RMPA) of 18 tests:
Proportion of Standardized Weighted Gaps (SWGs)
above 99th percentile (Chebyshev inequality)

First EV.	P			
	10	25	50	Mean
R				
0	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000
Mean	0.000	0.000	0.000	0.000

Second EV.	P			
	10	25	50	Mean
R				
0	0.006	0.000	0.013	0.006
5	0.006	0.006	0.000	0.004
7	0.000	0.000	0.000	0.000
Mean	0.004	0.002	0.004	0.003

Third EV.	P			
	10	25	50	Mean
R				
0	0.013	0.000	0.019	0.011
5	0.032	0.000	0.000	0.011
7	0.013	0.000	0.000	0.004
Mean	0.019	0.000	0.006	0.009

Distribution List

DR. TERRY ACKERMAN
EDUCATIONAL PSYCHOLOGY
260C EDUCATION BLDG.
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL 61801

DR. TERRY ALLARD
CODE 1142CS
OFFICE OF NAVAL RESEARCH
800 N. QUINCY ST.
ARLINGTON, VA 22217-5660

DR. NANCY ALLEN
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. GREGORY ANRIG
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. PHIPPS ARABIE
GRADUATE SCHOOL OF MANAGEMENT
RUTGERS UNIVERSITY
92 NEW STREET
NEWARK, NJ 07102-1895

DR. ISAAC I. BEJAR
LAW SCHOOL ADMISSIONS SERVICES
BOX 40
NEWTOWN, PA 18940-0040

DR. WILLIAM O. BERRY
DIRECTOR OF LIFE AND
ENVIRONMENTAL SCIENCES
AFOSR/NL, N1, BLDG. 410
BOLLING AFB, DC 20332-6448

DR. THOMAS G. BEVER
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF ROCHESTER
RIVER STATION
ROCHESTER, NY 14627

DR. MENUCHA BIRENBAUM
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. BRUCE BLOXOM
DEFENSE MANPOWER DATA CENTER
99 PACIFIC ST. SUITE 155A
MONTEREY, CA 93943-3231

DR. GWYNETH BOODOO
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. RICHARD L. BRANCH
HQ, USMEPCOM/MEPCT
2500 GREEN BAY ROAD
NORTH CHICAGO, IL 60064

DR. ROBERT BRENNAN
AMERICAN COLLEGE TESTING
PROGRAMS
P. O. BOX 168
IOWA CITY, IA 52243

DR. DAVID V. BUDESCU
DEPARTMENT OF PSYCHOLOGY
UNIV. OF IL, URBANA-CHAMPAIGN
603 E. DANIEL ST.
CHAMPAIGN, IL 61820

DR. GREGORY CANDELL
CTB/MACMILLAN/MCGRAW-HILL
2500 GARDEN ROAD
MONTEREY, CA 93940

DR. PAUL R. CHATELIER
PERCEPTRONICS
1911 NORTH FT. MYER DR.
SUITE LL00
ARLINGTON, VA 22209

DR. SUSAN CHIPMAN
COGNITIVE SCIENCE PROGRAM
OFFICE OF NAVAL RESEARCH
800 NORTH QUINCY ST.
ARLINGTON, VA 22217-5660

DR. RAYMOND E. CHRISTAL
UES LAMP SCIENCE ADVISOR AL/
HRMIL
BROOKS AFB, TX 78235

DR. NORMAN CLIFF
DEPARTMENT OF PSYCHOLOGY
UNIV. OF SO. CALIFORNIA
LOS ANGELES, CA 90089-1061

DIRECTOR
LIFE SCIENCES, CODE 1142
OFFICE OF NAVAL RESEARCH
ARLINGTON, VA 22217-5000

COMMANDING OFFICER
NAVAL RESEARCH LABORATORY
CODE 4827
WASHINGTON, DC 20375-5000

DR. JOHN M. CORNWELL
DEPARTMENT OF PSYCHOLOGY
I/O PSYCHOLOGY PROGRAM
TULANE UNIVERSITY
NEW ORLEANS, LA 70118

DR. WILLIAM CRANO
DEPARTMENT OF PSYCHOLOGY
TEXAS A&M UNIVERSITY
COLLEGE STATION, TX 77843

DR. LINDA CURRAN
DEFENSE MANPOWER DATA CENTER
SUITE 400
1600 WILSON BLVD
ROSSLYN, VA 22209

DR. TIMOTHY DAVEY
AMERICAN COLLEGE TESTING
PROGRAM
P.O. BOX 168
IOWA CITY, IA 52243

DR. CHARLES E. DAVIS
EDUCATIONAL TESTING SERVICE
MAIL STOP 22-T
PRINCETON, NJ 08541

DR. RALPH J. DEAYALA
MEASUREMENT, STATISTICS, AND
EVALUATION
BENJAMIN BLDG., RM. 1230F
UNIVERSITY OF MARYLAND
COLLEGE PARK, MD 20742

DR. SHARON DERRY
FLORIDA STATE UNIVERSITY
DEPARTMENT OF PSYCHOLOGY
TALLAHASSEE, FL 32306

HEI-KI DONG
BELLCORE
6 CORPORATE PL.
RM: PYA-1K207
P.O. BOX 1320
PISCATAWAY, NJ 08855-1320

DR. NEIL DORANS
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. FRITZ DRASGOW
UNIVERSITY OF ILLINOIS
DEPARTMENT OF PSYCHOLOGY
603 E. DANIEL ST.
CHAMPAIGN, IL 61820

DEFENSE TECHNICAL INFORMATION
CENTER
CAMERON STATION, BLDG 5
ALEXANDRIA, VA 22314
(2 COPIES)

DR. RICHARD DURAN
GRADUATE SCHOOL OF EDUCATION
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CA 93106

DR. SUSAN EMBRETSON
UNIVERSITY OF KANSAS
PSYCHOLOGY DEPARTMENT
426 FRASER
LAWRENCE, KS 66045

DR. GEORGE ENGELHARD, JR.
DIVISION OF EDUCATIONAL STUDIES
EMORY UNIVERSITY
210 FISHBURNE BLDG.
ATLANTA, GA 30322

ERIC FACILITY-ACQUISITIONS
2440 RESEARCH BLVD., SUITE 550
ROCKVILLE, MD 20850-3238

DR. MARSHALL J. FARR
FARR-SIGHT CO.
2520 NORTH VERNON STREET
ARLINGTON, VA 22207

DR. LEONARD FELDT
LINDQUIST CENTER FOR
MEASUREMENT
UNIVERSITY OF IOWA
IOWA CITY, IA 52242

DR. RICHARD L. FERGUSON
AMERICAN COLLEGE TESTING
P.O. BOX 168
IOWA CITY, IA 52243

DR. GERHARD FISCHER
LIEBIGGASSE 5
A 1010 VIENNA
AUSTRIA

DR. MYRON FISCHL
U.S. ARMY HEADQUARTERS
DAPE-HR
THE PENTAGON
WASHINGTON, DC 20310-0300

MR. PAUL FOLEY
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152-6800

CHAIR, DEPARTMENT OF COMPUTER SCIENCE
GEORGE MASON UNIVERSITY
FAIRFAX, VA 22030

DR. ROBERT D. GIBBONS
UNIVERSITY OF ILLINOIS AT CHICAGO
NPI 909A, M/C 913
912 SOUTH WOOD STREET
CHICAGO, IL 60612

DR. JANICE GIFFORD
UNIVERSITY OF MASSACHUSETTS
SCHOOL OF EDUCATION
AMHERST, MA 01003

DR. ROBERT GLASER
LEARNING RESEARCH & DEVELOPMENT CENTER
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15260

DR. SUSAN R. GOLDMAN
PEABODY COLLEGE, BOX 45
VANDERBILT UNIVERSITY
NASHVILLE, TN 37203

DR. TIMOTHY GOLDSMITH
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF NEW MEXICO
ALBUQUERQUE, NM 87131

DR. SHERRIE GOTT
AFHRL/MOMJ
BROOKS AFB, TX 78235-5601

DR. BERT GREEN
JOHNS HOPKINS UNIVERSITY
DEPARTMENT OF PSYCHOLOGY
CHARLES & 34TH STREET
BALTIMORE, MD 21218

PROF. EDWARD HAERTEL
SCHOOL OF EDUCATION
STANFORD UNIVERSITY
STANFORD, CA 94305-3096

DR. RONALD K. HAMBLETON
UNIVERSITY OF MASSACHUSETTS
LABORATORY OF PSYCHOMETRIC AND EVALUATIVE RESEARCH
HILLS SOUTH, ROOM 152
AMHERST, MA 01003

DR. DELWYN HARNISCH
UNIVERSITY OF ILLINOIS
51 GERTY DRIVE
CHAMPAIGN, IL 61820

DR. PATRICK R. HARRISON
COMPUTER SCIENCE DEPARTMENT
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402-5002

MS. REBECCA HETTER
NAVY PERSONNEL R&D CENTER
CODE 13
SAN DIEGO, CA 92152-6800

DR. THOMAS M. HIRSCH
ACT
P. O. BOX 168
IOWA CITY, IA 52243

DR. PAUL W. HOLLAND
EDUCATIONAL TESTING SERVICE, 21-T
ROSEDALE ROAD
PRINCETON, NJ 08541

PROF. LUTZ F. HORNKE
INSTITUT FUR PSYCHOLOGIE
RWTH AACHEN
JAEGERSTRASSE 17/19
D-5100 AACHEN
WEST GERMANY

MS. JULIA S. HOUGH
CAMBRIDGE UNIVERSITY PRESS
40 WEST 20TH STREET
NEW YORK, NY 10011

DR. WILLIAM HOWELL
CHIEF SCIENTIST
AFHRL/CA
BROOKS AFB, TX 78235-5601

DR. HUYNH HUYNH
COLLEGE OF EDUCATION
UNIV. OF SOUTH CAROLINA
COLUMBIA, SC 29208

DR. MARTIN J. IPPEL
CENTER FOR THE STUDY OF EDUCATION AND INSTRUCTION
LEIDEN UNIVERSITY
P. O. BOX 9555
2300 RB LEIDEN
THE NETHERLANDS

DR. ROBERT JANNARONE
ELEC. AND COMPUTER ENG. DEPT.
UNIVERSITY OF SOUTH CAROLINA
COLUMBIA, SC 29208

DR. KUMAR JOAG-DEV
UNIVERSITY OF ILLINOIS
DEPARTMENT OF STATISTICS
101 ILLINI HALL
725 SOUTH WRIGHT STREET
CHAMPAIGN, IL 61820

PROFESSOR DOUGLAS H. JONES
GRADUATE SCHOOL OF MANAGEMENT
RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY
NEWARK, NJ 07102

DR. BRIAN JUNKER
CARNEGIE-MELLON UNIVERSITY
DEPARTMENT OF STATISTICS
PITTSBURGH, PA 15213

DR. MARCEL JUST
CARNEGIE-MELLON UNIVERSITY
DEPARTMENT OF PSYCHOLOGY
SCHENLEY PARK
PITTSBURGH, PA 15213

DR. J. L. KAIWI
CODE 442/JK
NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CA 92152-5000

DR. MICHAEL KAPLAN
OFFICE OF BASIC RESEARCH
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333-5600

DR. JEREMY KILPATRICK
DEPARTMENT OF MATHEMATICS EDUCATION
105 ADERHOLD HALL
UNIVERSITY OF GEORGIA
ATHENS, GA 30602

MS. HAE-RIM KIM
UNIVERSITY OF ILLINOIS
DEPARTMENT OF STATISTICS
101 ILLINI HALL
725 SOUTH WRIGHT ST.
CHAMPAIGN, IL 61820

DR. JWA-KEUN KIM
DEPARTMENT OF PSYCHOLOGY
MIDDLE TENNESSEE STATE UNIVERSITY
MURFREESBORO, TN 37132

DR. SUNG-HOON KIM KEDI
92-6 UMYEON-DONG
SEOCHO-GU SEOUL
SOUTH KOREA

DR. G. GAGE KINGSBURY
PORTLAND PUBLIC SCHOOLS
RESEARCH AND EVALUATION DEPARTMENT
501 NORTH DIXON STREET
P. O. BOX 3107
PORTLAND, OR 97209-3107

DR. WILLIAM KOCH
BOX 7246, MEAS. AND EVAL. CTR.
UNIVERSITY OF TEXAS-AUSTIN
AUSTIN, TX 78703

DR. JAMES KRAATZ
COMPUTER-BASED EDUCATION RESEARCH LABORATORY
UNIVERSITY OF ILLINOIS
URBANA, IL 61801

DR. PATRICK KYLLONEN
AFHRL/MOEL
BROOKS AFB, TX 78235

MS. CAROLYN LANEY
1515 SPENCERVILLE ROD
SPENCERVILLE, MD 20868

RICHARD LANTERMAN
COMMANDANT (G-PWP)
US COAST GUARD
2100 SECOND ST., SW
WASHINGTON, DC 20593-0001

DR. MICHAEL LEVINE
EDUCATIONAL PSYCHOLOGY
210 EDUCATION BLDG.
1310 SOUTH SIXTH STREET
UNIVERSITY OF IL AT URBANA-CHAMPAIGN
CHAMPAIGN, IL 61820-6990

DR. CHARLES LEWIS
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541-0001

MR. HSIN-HUNG LI
UNIVERSITY OF ILLINOIS
DEPARTMENT OF STATISTICS
101 ILLINI HALL
725 SOUTH WRIGHT ST.
CHAMPAIGN, IL 61820

LIBRARY
NAVAL TRAINING SYSTEMS CENTER
12350 RESEARCH PARKWAY
ORLANDO, FL 32826-3224

DR. MARCIA C. LINN
GRADUATE SCHOOL OF EDUCATION,
EMST TOLMAN HALL
UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720

DR. ROBERT L. LINN
CAMPUS BOX 249
UNIVERSITY OF COLORADO
BOULDER, CO 80309-0249

LOGICON INC. (ATTN: LIBRARY)
TACTICAL AND TRAINING SYSTEMS
DIVISION
P.O. BOX 85158
SAN DIEGO, CA 92138-5158

DR. RICHARD LUECHT
ACT
P. O. BOX 168
IOWA CITY, IA 52243

DR. GEORGE B. MACREADY
DEPARTMENT OF MEASUREMENT
STATISTICS & EVALUATION
COLLEGE OF EDUCATION
UNIVERSITY OF MARYLAND
COLLEGE PARK, MD 20742

DR. EVANS MANDES
GEORGE MASON UNIVERSITY
4400 UNIVERSITY DRIVE
FAIRFAX, VA 22030

DR. PAUL MAYBERRY
CENTER FOR NAVAL ANALYSIS
4401 FORD AVENUE
P.O. BOX 16268
ALEXANDRIA, VA 22302-0268

DR. JAMES R. MCBRIDE
HUMRRO
6430 ELMHURST DRIVE
SAN DIEGO, CA 92120

MR. CHRISTOPHER MCCUSKER
UNIVERSITY OF ILLINOIS
DEPARTMENT OF PSYCHOLOGY
603 E. DANIEL ST.
CHAMPAIGN, IL 61820

DR. ROBERT MCKINLEY
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. JOSEPH MCLACHLAN
NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
CODE 14
SAN DIEGO, CA 92152-6800

ALAN MEAD
C/O DR. MICHAEL LEVINE
EDUCATIONAL PSYCHOLOGY
210 EDUCATION BLDG.
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL 61801

DR. TIMOTHY MILLER
ACT
P. O. BOX 168
IOWA CITY, IA 52243

DR. ROBERT MISLEVY
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. IVO MOLENAR
FACULTEIT SOCIALE
WETENSCHAPPEN
RIJKSUNIVERSITEIT GRONINGEN
GROTE KRUISSTRAAT 2/1
9712 TS GRONINGEN
THE NETHERLANDS

DR. E. MURAKI
EDUCATIONAL TESTING SERVICE
ROSEDALE ROAD
PRINCETON, NJ 08541

DR. RATNA NANDAKUMAR
EDUCATIONAL STUDIES
WILLARD HALL, ROOM 213E
UNIVERSITY OF DELAWARE
NEWARK, DE 19716

ACADEMIC PROGS. & RESEARCH
BRANCH
NAVAL TECHNICAL TRAINING
COMMAND
CODE N-62
NAS MEMPHIS (75)
MILLINGTON, TN 30854

DR. W. ALAN NICEWANDER
UNIVERSITY OF OKLAHOMA
DEPARTMENT OF PSYCHOLOGY
NORMAN, OK 73071

HEAD, PERSONNEL SYSTEMS
DEPARTMENT
NPRDC (CODE 12)
SAN DIEGO, CA 92152-6800

DIRECTOR
TRAINING SYSTEMS DEPARTMENT
NPRDC (CODE 14)
SAN DIEGO, CA 92152-6800

LIBRARY, NPRDC
CODE 041
SAN DIEGO, CA 92152-6800

LIBRARIAN
NAVAL CENTER FOR APPLIED
RESEARCH IN ARTIFICIAL
INTELLIGENCE
NAVAL RESEARCH LABORATORY
CODE 5510
WASHINGTON, DC 20375-5000

DEPT. OF THE NAVY
ONR RESIDENT REPRESENTATIVE
MASSACHUSETTS INSTITUTE OF
TECHNOLOGY
495 SUMMER STREET, ROOM 103
BOSTON, MA 02210-2109

OFFICE OF NAVAL RESEARCH
CODE 3422
800 N. QUINCY STREET
ARLINGTON, VA 22217-5660
(6 COPIES)

SPECIAL ASSISTANT FOR RESEARCH
MANAGEMENT
CHIEF OF NAVAL PERSONNEL (PERS-
01JT)
DEPARTMENT OF THE NAVY
WASHINGTON, DC 20350-2000

DR. JUDITH ORASANU
MAIL STOP 239-1
NASA AMES RESEARCH CENTER
MOFFETT FIELD, CA 94035

DR. PETER J. PASHLEY
EDUCATIONAL TESTING SERVICE
ROSEDALE ROAD
PRINCETON, NJ 08541

WAYNE M. PATIENCE
AMERICAN COUNCIL ON EDUCATION
GED TESTING SERVICE, SUITE 20
ONE DUPONT CIRCLE, NW
WASHINGTON, DC 20036

DEPT. OF ADMINISTRATIVE
SCIENCES CODE 54
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5026

DR. PETER PIROLLI
SCHOOL OF EDUCATION
UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720

DR. MARK D. RECKASE
ACT
P. O. BOX 168
IOWA CITY, IA 52243

MR. STEVE REISE
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF CALIFORNIA
RIVERSIDE, CA 92521

MR. LOUIS ROUSSOS
UNIVERSITY OF ILLINOIS
DEPARTMENT OF STATISTICS
101 ILLINI HALL
725 SOUTH WRIGHT ST.
CHAMPAIGN, IL 61820

DR. DONALD RUBIN
STATISTICS DEPARTMENT
SCIENCE CENTER, ROOM 608
1 OXFORD STREET
HARVARD UNIVERSITY
CAMBRIDGE, MA 02138

DR. FUMIKO SAMEJIMA
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
310B AUSTIN PEAY BLDG.
KNOXVILLE, TN 37966-0900

MR. DREW SANDS
NPRDC CODE 62
SAN DIEGO, CA 92152-7250

DR. MARY SCHRATZ
4100 PARKSIDE
CARLSBAD, CA 92008

MR. ROBERT SEMMES
N218 ELLIOTT HALL
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN 55455-0344

DR. VALERIE L. SHALIN
DEPARTMENT OF INDUSTRIAL ENGINEERING
STATE UNIVERSITY OF NEW YORK
342 LAWRENCE D. BELL HALL
BUFFALO, NY 14260

MR. RICHARD J. SHAVELSON
GRADUATE SCHOOL OF EDUCATION
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CA 93106

MS. KATHLEEN SHEEHAN
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. KAZUO SHIGEMASU
7-9-24 KUGENUMA-KAIGAN
FUJISAWA 251
JAPAN

DR. RANDALL SHUMAKER
NAVAL RESEARCH LABORATORY
CODE 5500
4555 OVERLOOK AVENUE, S.W.
WASHINGTON, DC 20375-5000

DR. JUDY SPRAY
ACT
P.O. BOX 168
IOWA CITY, IA 52243

DR. MARTHA STOCKING
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. WILLIAM STOUT
UNIVERSITY OF ILLINOIS
DEPARTMENT OF STATISTICS
101 ILLINI HALL
725 SOUTH WRIGHT ST.
CHAMPAIGN, IL 61820

DR. KIKUMI TATSUOKA
EDUCATIONAL TESTING SERVICE
MAIL STOP 03-T
PRINCETON, NJ 08541

DR. DAVID THISSEN
PSYCHOMETRIC LABORATORY
CB# 3270, DAVIE HALL
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NC 27599-3270

MR. THOMAS J. THOMAS
FEDERAL EXPRESS CORPORATION
HUMAN RESOURCE DEVELOPMENT
3035 DIRECTOR ROW, SUITE 501
MEMPHIS, TN 38131

MR. GARY THOMASSON
UNIVERSITY OF ILLINOIS
EDUCATIONAL PSYCHOLOGY
CHAMPAIGN, IL 61820

DR. HOWARD WAINER
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

ELIZABETH WALD
OFFICE OF NAVAL TECHNOLOGY
CODE 227
800 NORTH QUINCY STREET
ARLINGTON, VA 22217-5000

DR. MICHAEL T. WALLER
UNIVERSITY OF WISCONSIN-MILWAUKEE
EDUCATIONAL PSYCHOLOGY DEPT.
BOX 413
MILWAUKEE, WI 53201

DR. MING-MEI WANG
EDUCATIONAL TESTING SERVICE
MAIL STOP 03-T
PRINCETON, NJ 08541

DR. THOMAS A. WARM
FAA ACADEMY
P.O. BOX 25082
OKLAHOMA CITY, OK 73125

DR. DAVID J. WEISS
N660 ELLIOTT HALL
UNIVERSITY OF MINNESOTA
75 E. RIVER ROAD
MINNEAPOLIS, MN 55455-0344

DR. DOUGLAS WETZEL
CODE 15
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152-6800

GERMAN MILITARY REPRESENTATIVE
PERSONALSTAMMAMT
KOELNER STR. 262
D-5000 KOELN 90
WEST GERMANY

DR. DAVID WILEY
SCHOOL OF EDUCATION AND SOCIAL POLICY
NORTHWESTERN UNIVERSITY
EVANSTON, IL 60208

DR. BRUCE WILLIAMS
DEPARTMENT OF EDUCATIONAL PSYCHOLOGY
UNIVERSITY OF ILLINOIS
URBANA, IL 61801

DR. MARK WILSON
SCHOOL OF EDUCATION
UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720

DR. EUGENE WINOGRAD
DEPARTMENT OF PSYCHOLOGY
EMORY UNIVERSITY
ATLANTA, GA 30322

DR. MARTIN F. WISKOFF
PERSEREC
99 PACIFIC ST., SUITE 4556
MONTEREY, CA 93940

MR. JOHN H. WOLFE
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152-6800

DR. KENTARO YAMAMOTO
03-0T
EDUCATIONAL TESTING SERVICE
ROSEDALE ROAD
PRINCETON, NJ 08541

MS. DUANLI YAN
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08541

DR. WENDY YEN
CTB/MCGRAW HILL
DEL MONTE RESEARCH PARK
MONTEREY, CA 93940

DR. JOSEPH L. YOUNG
NATIONAL SCIENCE FOUNDATION
ROOM 320
1800 G STREET, N.W.
WASHINGTON, DC 20550