AD-A143 232

THE STATISTICAL ASSESSMENT OF LATENT TRAIT DIMENSIONALITY IN PSYCHOLOGICAL TESTING

FINAL REPORT 8-82-ONR

WILLIAM F. STOUT

Department of Mathematics
University of Illinois
at Urbana-Champaign

June 1984

This research was sponsored by the
Personnel and Training Research Programs,
Psychological Science Division, Office of Naval Research,
under Contract No. N00014-82-K-0486; NR 150-488.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Final Report 8-82-ONR	2. GOVT ACCESSION NO. AD-A243 232	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Statistical Assessment of Latent Trait Dimensionality in Psychological Testing.		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William Stout		8. CONTRACT OR GRANT NUMBER(s) N00014-82-K-0436
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics, University of Illinois 1409 West Green Street Urbana, IL 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N; RR042-04; RR 042-04-01 NR 150-488
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 442PT) 800 N. Quincy St., Arlington, VA 22217		12. REPORT DATE May, 1984
		13. NUMBER OF PAGES 60
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Item response theory, latent trait theory, dimensionality, unidimensionality, statistical test of unidimensionality, large sample theory

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Assuming a nonparametric item response theory model, a large sample procedure for testing the unidimensionality of the latent ability space is proposed. Under the assumption of unidimensionality, the asymptotic distribution of the test statistic is derived, thereby establishing an asymptotically valid statistical test of unidimensionality. A rigorous mathematical definition of dimensionality is proposed as an alternative to the classical item response theory definition of dimensionality. This new definition, while item response theory based, is more analogous to the factor analytic notion of dimensionality

DD FORM 1 JAN 73 1473

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) and as such is more congruent with the conception of dimensionality held by applicators. The statistical test procedure is shown to have asymptotic power 1 for multidimensional (in our sense) tests. Monte Carlo studies, using a specially written Fortran program, indicate good agreement with the nominal level of significance when unidimensionality holds and good power performance for examinee sample sizes and psychological test lengths often encountered in practice. This adherence to the nominal level of significance and good power performance is also the case, as desired, when the items are "multiple determined", thereby producing a classical dimensionality much greater than 1. Finally, an analogous test procedure is proposed for testing $H: d \leq d_0$ vs. $A: d > d_0$ for fixed $d_0 \geq 2$, d denoting dimensionality.

June, 1984

Note concerning the relationship of enclosed final report to recently distributed technical report

You almost certainly recently received an ONR technical report entitled "A Statistical Procedure for Assessing Test Dimensionality," which I wrote in the fall of 1981. The enclosed Final Report totally supersedes the 1981 document, which was an effort to articulate my early work on the dimensionality problem. Hence, you should concentrate your time on the enclosed Final Report, which carefully describes the current status of my dimensionality research in its appropriate psychometric context.

William Stout

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

The Statistical Assessment of Latent Trait Dimensionality in Psychological Testing

(Final Technical Report of N00014-82-K-0486)

William Stout
University of Illinois at Urbana-Champaign

1. Introduction

In Lord (1980), Frederick Lord states "There is a great need for a statistical significance test for the unidimensionality of a set of test items." Indeed, it is an important problem to be able to determine whether a test that purports to measure the level of a certain ability is in reality significantly contaminated by the varying levels of one of more other abilities displayed by examinees taking the test. For example, is a test of mathematical ability contaminated by varying levels of verbal ability displayed by subjects taking the test? As a second example, it is of serious concern in Canada that performance on standardized ability tests not be influenced by varying examinee of familiarity with the French Canadian or Anglo-Canadian cluture (see Sarrazin (198 Further, the standard item response theory (IRT) methodology (e.g., LOGIST) is predicated upon the assumption of unidimensionality. Because of the large number of private and governmental organizations routinely using tests to screen people by assessing their levels of various aptitudes or abilities, the problem of assessing the dimensionality (especially the unidimensionality) of a test is of great importance. Now, with the imminent use of computerized adaptive testing by the U. S. Armed Forces (a test setting for which any two examinee taking the same test will in general be administered non-identical sets of items the issue of dimensionality becomes particularly critical.

The basic objective of the project was to develop a practicable, theoretically based statistical procedure for testing for the unidimensionality of a test. To this end, the following objectives have been attained:

(1) The development of an intuitively plausible statistical method for testing for unidimensionality.

(2) The (rather delicate) derivation of the asymptotic distribution of the proposed test statistic (of (1)) under the assumptions of unidimensionality and reasonable regularity conditions, thereby establishing an asymptotic level α test of unidimensionality.

(3) The development of a rigorous mathematical definition of test dimensionality that is item response theory based but yet is consistent with the classical factor analytic notion of dimensionality and also is not adversely influenced by the inherent multidimensionality of individual test items.

(4) The establishment of asymptotic power one for the statistical test whenever unidimensionality in the sense of the definition referred to in (3) fails to hold.

(5) The writing of a FORTRAN program to simulate the use of the statistical test procedure for the case of a unidimensional test and for the case of a two dimensional test. Here the two dimensional test was modeled both with and without multiply determined (i.e., multidimensional) test items.

(6) The conducting of preliminary Monte Carlo studies to assess the practicibility of the proposed test of unidimensionality for realistic test lengths and examinee population sizes, these results indicating reasonable faithfulness to the prescribed level of significance, even in the presence of multidimensiona items, and indicating reasonable power.

(7) The development of an analogous test to that of unidimensionality in order to test (letting d denote dimensionality)

$$H_0 : d \leq d_0 \quad \text{vs} \quad H_1 : d > d_0$$

where d_0 is a fixed integer satisfying $d_0 \geq 2$.

The remainder of the report is devoted to a description of the attainment of these objectives together with certain prerequisite background information.

2. A careful statement of the item response theory model.

The mathematical definition of dimensionality and the mathematical derivation of the large sample theory for the proposed test statistic both require a careful statement of the assumptions of the item response theory model:

Consider sampling J examinees from a population and administering a test consisting of N items to each sampled examinee. Suppose each item is scored correct or incorrect. (1 for correct and 0 for incorrect).

Suppose that associated with each examinee is a vector-valued variable $\underline{\theta} \equiv (\theta_1, \theta_2, \ldots, \theta_d)$ (the "ability" of the examinee) that determines the probability of a correct response for each item administered to the examinee. For each sampled examinee j, binary random variables $\{U_{ij}\}$ are observed, where i is the item index. The $\{U_{ij}\}$ are the observed or manifest variables referred to in the discussion of latent structure models. "Item characteristic curves" are defined by

$$(2.1) \qquad P_i(\underline{\theta}) = P[U_{ij} = 1 \mid \underline{\Theta}_j = \underline{\theta}] = 1 - P[U_{ij} = 0 \mid \underline{\Theta}_j = \underline{\theta}],$$

the probability of a correct response to item i, given that the jth sampled examinee has ability $\underline{\theta}$. Here it is assumed that examinees are randomly sampled from an infinite population, thereby inducing a probability distribution on the $\underline{\theta}$ ability space with associated random vector $\underline{\Theta}$. Here we let $(\underline{\Theta}_1, \underline{\Theta}_2, \ldots, \underline{\Theta}_J)$ denote the random vector of abilities for J sampled examinees. Throughout, $\underline{\Theta}$ is assumed to be a continuous random vector with density denoted by $f(\underline{\theta})$. L denotes the domain of $f(\underline{\theta})$ and is a subset of d dimensional Euclidean space.

Item response theory is based upon the fundamental assumption of local independence, which is assumed throughout this report. The intuitive idea is that, conditional on knowing the ability of an examinee, the item responses are independent of each other. That is, for example, the dependence between U_{i1} and $U_{i'1}$ for $i \neq i'$ is entirely explained ("mediated") by knowing the value of $\underline{\Theta}_1$. Formally, local independence is said to hold provided

$$(2.2) \qquad P[U_{11} = u_1, U_{21} = u_2, \ldots, U_{N1} = u_N \mid \underline{\Theta}_1 = \underline{\theta}] = \prod_{i=1}^{N} P[U_{i1} = u_i \mid \underline{\Theta}_1 = \underline{\theta}]$$

for all $\underline{\theta} \in L$ and each choice of $u_i = 0$ or 1. Recall that N denotes the total number of items.

It is also necessary to assume throughout two assumptions about the sampling of examinees and the independence of information provided by different examinees. It is assumed throughout that

$$(2.3) \qquad (\underline{\Theta}_1, \underline{\Theta}_2, \ldots, \underline{\Theta}_J)$$

consists of independent identically distributed (iid) random vectors. That is, the $\{\underline{\Theta}_j\}$ are assumed to consist of a random sample.

Now let $\underline{U}_j = (U_{1j}, U_{2j}, \ldots, U_{Nj})$, $\underline{U}_j$ denoting the test performance. It is assumed that knowing each examinee's $\underline{\theta}$ produces a sort of conditional independence: That is, it is assumed throughout that

$$(2.4) \qquad P[\underline{U}_1 = \underline{u}_1, \underline{U}_2 = \underline{u}_2, \ldots, \underline{U}_J = \underline{u}_J \mid \underline{\Theta}_1 = \underline{\theta}_1, \underline{\Theta}_2 = \underline{\theta}_2, \ldots, \underline{\Theta}_J = \underline{\theta}_J]$$
$$= \prod_{i=1}^{J} P[\underline{U}_j = \underline{u}_j \mid \underline{\Theta}_j = \underline{\theta}_j]$$

In other words, the responses of different units are, conditional on knowing their latent values, independent of each other.

For some purposes the classical definition of dimensionality used in item response theory will be used in this report (see Lord and Novick (1968), pp. 359-362 for a good discussion of this). This amounts to stating that d is the correct dimensionality provided the joint distribution of the observed random variables $\{U_{ij}, 1 \leq i \leq N, 1 \leq j \leq J\}$ is representable in terms of a d dimensional distribution of i.i.d. $\underline{\Theta}_j$ and identical (in j) conditional distributions of $\underline{U}_j$, given $\underline{\Theta}_j$, with local independence and (2.4) assumed, and moreover that such a representation is impossible for $d' < d$. Here of course by $\underline{\Theta}$ being d dimensional is simply meant (since $\underline{\Theta}$ is assumed to be a random vector of continuous type) that the density $f(\theta)$ of $\underline{\Theta}$ has d dimensional Euclidean space as its domain.

It will be assumed throughout the report, that, denoting here the classical dimensionality by d, whenever the joint distribution of the $\{U_{ij}\}$ is modeled in terms of (identically distributed) latent variables $\{\underline{\Theta}_j\}$ that the distribution of $\underline{\Theta}$ is in fact d dimensional. Such a remark is necessary, since, of course, it is always possible to use a distribution for $\underline{\Theta}$ that has dimensionality larger than d.

3. A statistical approach to the assessment of latent structure unidimensionality.

As stated in Section 1, the main object of the project was to develop a statistical method for assessing the choice of dimensionality d. The dimensionality test of most importance in mental test theory and likely in most other settings too is the test of unidimensionality:

(3.1) $\quad H : d = 1 \quad \text{vs} \quad A : d > 1.$

Our hypothesis testing procedure to investigate (3.1) is now broken into steps. By "score" is always meant proportion correct.

Steps of the procedure:

Step 1: Split the N test items into a long subtest of length n and a short subtest of length M. These two subtests are called respectively: the partitioning subtest (length n) and the assessment subtest (length M). Observations taken: $\{U_{ij} : i = 1,2,\ldots,n+M;\ j = 1,2,\ldots,J\}$. (Typical values in mental test applications: $J = 1000$, $N = 75$).

Step 2: Partition the unit interval $[0,1]$ into subintervals $\{A_k^{(n)}, 1 \leq k \leq K_n\}$ of equal length. E.g., typical interval: $[0.48, 0.52)$, in which case there are 25 intervals.

Step 3: Assign examinees to partition subintervals according to their scores on the partitioning subtest. E.g., if an examinee's partitioning subtest score is 0.51 and if 0.04 is the interval length then the examinee is assigned to the interval $[0.48, 0.52)$. Let $J_k^{(n)}$ be the number of examinees assigned to the kth interval of the nth partition, the "nth partition" being the partition associated with a partitioning subtest of length n.

Step 4: (Construction of the test statistic). Resubscript the U_{ij} for $n+1 \le i \le n+M$ such that $U_{ijk} \equiv U^{(n)}_{ijk}$ indicates the correctness of the response of examinee j of subinterval k of partition n $(1 \le j \le J^{(n)}_k)$ to item i. Note that $J = \sum_{k=1}^{K_n} J^{(n)}_k$. In each subinterval of the partition that contains enough examinees (according to some convention set by the statistical user), the difference of two variance estimates is computed. From this point on, K_n will denote the number of such intervals rather than the number of intervals of the partition and $A^{(n)}_1, \ldots, A^{(n)}_{K_n}$ will denote the subintervals containing enough examinees. Computation of $\hat{\sigma}^2_{Y,k}$, the first variance estimate for interval k: Let

$$Y^{(k)}_j = \Sigma_{i=n+1}^{n+m} U_{ijk}/M$$

the assessment subtest score of the jth examinee of interval k on the assessment subtest. (E.g., $j = 2$ denotes the second examinee among the $J^{(n)}_k$ examinees assigned to interval k of the nth partition.) Let

$$\bar{Y}^{(k)} = \Sigma_{j=1}^{J^{(n)}_k} Y^{(k)}_j / J^{(n)}_k, \tag{3.2}$$

the average of examinee assessment subtest scores for interval k. Let

$$\hat{\sigma}^2_{Y,k} = \Sigma_{j=1}^{J^{(n)}_k} (Y^{(k)}_j - \bar{Y}^{(k)})^2 / J^{(n)}_k. \tag{3.3}$$

Computation of $\sigma^2_{P,k}$, the second variance estimate for interval k:

Let

$$\hat{p}^{(k)}_i = \Sigma_{j=1}^{J^{(n)}_k} U_{ijk}/J^{(n)}_k \tag{3.4}$$

$$\hat{\sigma}^2_{P,k} = \Sigma_{i=n+1}^{n+M} \hat{p}^{(k)}_i (1 - \hat{p}^{(k)}_i)/M^2. \tag{3.5}$$

The test procedure is then to reject H if

$$\sum_{k=1}^{K_n} (\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}),$$

appropriately normalized, is sufficiently large.

Normalization: Let

$$(3.6)\quad \begin{cases} \hat{\mu}_4(Y) = \sum_{j=1}^{J_k^{(n)}} (Y_j^{(k)} - \bar{Y}^{(k)})^4 / J_k^{(n)} \\ \\ \hat{\delta}_{4,k} = \sum_{i=1}^{M} \hat{p}_i^{(k)} (1 - \hat{p}_i^{(k)}) (1 - 2\hat{p}_i^{(k)})^2 \end{cases}$$

Finally, let

$$(3.7)^*\quad T \equiv T_n = \frac{\sum_{k=1}^{K_n} (\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k})}{\left\{\sum_{k=1}^{K_n} \left[(\hat{\mu}_4(Y) - \hat{\sigma}^4_{Y,k}) - \frac{\hat{\delta}_{4,k}}{M^4}\right] / J_k^{(n)}\right\}^{1/2}}$$

Then, the procedure is to reject the null hypothesis of (4.1) provided

$$(3.8)\quad T > Z_\alpha,$$

where Z_α is the upper $100(1-\alpha)$ percentile value for a standard normal distribution.

Remark. The selection of which items are assigned to the partitioning subtest and which items will be assigned to the assessment subtest is related to the power of the test and will be discussed in Section 8.

*See page 30 for a modification of T that may have better small sample properties.

It seems useful to have an intuitive understanding of the procedure given by (3.8). For this purpose let $K_n = 1$ and ignore the scaling provided by the denominator in (3.7). That is, think of the statistical procedure as rejecting the hypothesis of unidimensionality when $\hat{\sigma}^2_{Y,1} - \hat{\sigma}^2_{P,1}$ is "large" $(k = 1)$. Recall from classical test theory that for a test of length M with Y denoting proportion right and α_m denoting the probability that a randomly selected examinee gets item m right, that the Kuder-Richardson formula-20 coefficient ρ_{20} is defined by (e.g., see Lord (1980), p. 8)

$$\rho_{20} = \frac{M}{M-1}\left[1 - \frac{\sum_{m=1}^{M} \alpha_m(1-\alpha_m)/M^2}{\sigma^2_Y}\right]$$

(the division by M^2 occurs because Y denotes proportion correct rather than number correct). Note that

(3.9) $\qquad \rho_{20} = 0$ if and only if $\sigma^2_Y - \sum_{m=1}^{M}\alpha_m(1-\alpha_m)/M^2 = 0.$

Of course $\rho_{20} = 0$ is to be psychometrically interpreted (up to an approximation) as the fact that Y has no reliability at all. Although unrelated to the author's process of discovery of his unidimensionality test, (3.9) can be nicely used to give a simple psychometric interpretation and justification for the test procedure (3.8): In the present test setting, $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$ is an estimate of $\sigma^2_Y - \sum_{m=1}^{M}\alpha_m(1-\alpha_m)/M^2$, m indexing the M items of the assessment subtest. Here α_m and σ^2_Y are computed for the "population" of examinees assigned to subinterval 1 by the assessment subtest. Thus $\hat{\sigma}^2_{Y,1} - \hat{\sigma}^2_{P,1}$ "large" is evidence that ρ_{20} is greater than 0. Thus, the procedure rejects the hypothesis H_0 of unidimensionality if and only if the assessment subtest for subinterval 1 examinees shows statistical evidence of having some reliability. That is, H_0 is rejected if and only if there is statistical evidence that the assessment subtest provides information about examinee "true score" (or latent ability from the item response theory viewpoint) beyond that provided by the knowledge of which partitioning subinterval an examinee is assigned to by the partitioning subtest. And, the assessment subtest can only provide such information (except for negligible finite sample error) provided there is more than one dimension being measured.

4. The large sample theory when $d = 1$.

In this section the classical conception of dimensionality described at the end of Section 2 will be used. The main purpose of this section is to present the precise conditions under which T, as defined by (3.7), can be shown to have asymptotically a standard normal (i.e., N(0,1)) distribution.

It is however easy to see several reasons why one might also wish to have the large sample distribution of the $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$, suitably standardized, for each subinterval k. First, in some applications K_n, the number of subintervals used in the construction of the test statistic may be rather small and hence the large sample theory for T fail to provide a good approximation for the actual distribution of T. In this case, having the large sample distribution for each $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$ in T then, using the independence of the $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$, yields a large sample theory for T when K_n is small. Second, one may wish to construct a new test statistic for unidimensionality still based on the summands of T, but with the new test statistic having a different formula from that of T. Finally, and perhaps most important, one may be interested in assessing the contribution to lack of dimensionality resulting from only a portion of the latent ability spac That is, one may wish to construct a test statistic based upon only certain subintervals of the partition, indeed possibly even using only one subinterval. Because of the above, it is desirable to obtain the large sample distribution of each $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$ as the number of test items and examinees each become suitably large. Convergence in law (i.e., in distribution) will be denoted throughout by $\xrightarrow{L}$. $N(\mu,\sigma^2)$ denotes a normal distribution with mean μ and variance σ^2. Z always denotes a N(0,1) random variable. For example, the statement that X_n converges in law to a standard normal random variable would be denoted by

$$X_n \xrightarrow{L} Z.$$

The two basic large sample theorems of this section are based on a set of assumptions, which are now stated, together with a brief description of the practical interpretation and reasonableness of each assumption.

First it is assumed that

$$d = 1 \tag{4.1}$$

in the classical sense of dimensionality. That is, the large sample distributions of the $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$ and T are stated in this section under the assumption that the hypothesis $H : d = 1$ is true, true in the classical sense of dimensionality.

It is assumed that

$$\min_{1 \leq k \leq K_n} J_k^{(n)} \to \infty \quad \text{as} \quad n \to \infty. \tag{4.2}$$

This simply amounts to excluding partition subintervals that contain too few examinees. In practice, some suitable convention would be used, such as $J_k^{(n)} \geq 20$ for each $1 \leq k \leq K_n$.

The number of items M on the assessment subtest is assumed to be fixed as n, the number of items on the partitioning subtest, approaches ∞. This means, in effect, that the number of items on the partitioning subtest must be small compared to the number of items on the assessment subtest. This amounts to a genuine restriction in the design of the statistical test in that our theory simply does not support taking $M = n = 38$ when $N = 76$ for example. However, if $M = 5$, $n = 71$ produces good results in terms of attaining a reasonably powerful test with the actual level of significance being close to the nominal level of significance, then the restriction becomes non-essential. And, indeed, preliminary Monte Carlo studies support the preceding statement. (See Section 7.)

It is assumed that the M assessment items are the same for every n. This is merely a virtual restriction, having no "real world" content since in fact in practice one applies the procedure for a particular set of M items and only one choice of n. Although a double subscript notation would avoid notational inconsistency, we nonetheless denote the item characteristic curves (ICC's) of these fixed M items by

$$(4.3) \qquad P_{n+1}(\theta), P_{n+2}(\theta), \ldots, P_{n+M}(\theta); \; -\infty < \theta < \infty.$$

Now, the ICC's for the items of the partitioning subtest are denoted by

$$(4.4) \qquad P_1(\theta), P_2(\theta), \ldots, P_n(\theta),$$

with the understanding that $P_i(\theta)$ does not necessarily denote the same ICC on the n'th partitioning subtest as it does on the nth partitioning subtest for $n \neq n'$. Although assuming that $P_i(\theta)$ is the same for all choices of partition subtest length $n \geq 1$ would only be a virtual rather than a practical restriction, nonetheless we do not need to make this restriction.

Certain assumptions are made about the form of the ICC's. First each ICC is assumed to be continuous and differentiable in θ and strictly increasing in Clearly this entirely reasonable restriction needs no explanation nor defense.

It is assumed that there exists an interval (a,b) such that for some likely small number ε satisfying $\frac{1}{4} > \varepsilon > 0$ that

$$(4.5) \qquad P_i'(\theta) \geq \varepsilon \quad \text{for } a \leq \theta \leq b,\ 1 \leq i \leq N,$$

$$(4.6) \qquad P_i'(\theta) \leq C \quad \text{for all } \theta,\ 1 \leq i \leq N, \text{ some (possibly large) } C, \text{ and}$$

$$(4.7) \qquad \varepsilon \leq P_i(\theta)(1 - P_i(\theta)) \leq 1 - \varepsilon \quad \text{for } a \leq \theta \leq b,\ 1 \leq i \leq N.$$

Here (4.5) merely states that there is some ability range over which items are uniformly (both as θ varies and items change) sensitive to ability differences. This merely amounts to discarding items, as test constructors do, that are not sufficiently discriminating or are too easy or too difficult. This practice of discarding overly easy and difficult items also justifies (4.7), which amounts to not allowing items that are too easy or too difficult. (4.6) of course states that, even locally in θ, none of the items are allowed to be too discriminating. It is of course impossible in virtually all test settings to constrict highly discriminating items, even if sometimes desirable to do so. As Lord and Novick (1968, p. 379) state, "We might note that it is rare to find values of a_g (slope) as large as 2 in aptitude and achievement testing."

Let $P_i(-\infty) = \lim_{\theta \to -\infty} P_i(\theta)$. Then let $\eta = \sup_{i \geq 1} P_i(-\infty)$. It is assumed that each subinterval of the partition included in the construction of T lies in the interval $(\eta, 1]$. This is merely a technical restriction that will cause no trouble in practice; and indeed, if needed, this restriction could be weakened. Typically, $\eta = 0$, $1/4$, or $1/5$ is assumed, based upon various assumptions about "guessing."

It is required that

$$(4.8) \qquad J_k^{(n)} K_n \leq Cn^2$$

for all k, n for some C in order to obtain the distribution of T. If one is merely interested in the distribution of each of the $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$ separately then (4.8) is weakened to

(4.9) $\quad J_k^{(n)} \le Cn^2$

for all k, n for some C. These two assumptions restrict the number of examinees per subinterval and in the case of the distribution of T, (4.8) restr the total number of intervals contributing to T. This is a genuine restriction, but, fortunately, allowing $J_k^{(n)} K_n$ to be the order of n^2 seems reasonable for most applications. E.g., if the partitioning subtest of a test consis of 100 items, then $J_n^{(k)} = 30$ for each k, $K_n = 30$ is such that $J_k^{(n)} K_n$ is "small" compared to $n^2 = 10^4$. Preliminary Monte Carlo studies again seem to support the accuracy of the large sample approximation of T for reasonable choices of n and $\Sigma_{k=1}^{K_n} J_k^{(n)}$, the number of test examinees (ignoring discarded intervals here).

The following assumption is needed for highly technical reasons. It could undoubtedly be removed, at the expense of a rather complicated answer for the large sample distributions of $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$ and T. It is assumed that there exists some (possibly small) $\varepsilon > 0$ such that for all k, n, and each of the M assessment items (indexed by i),

(4.10) $\quad |P_i(\theta_k^{(n)}) - \tfrac{1}{2}| > \varepsilon$

where $\theta_k^{(n)}$ is defined by, letting Y denote proportion correct on the partitioning subtest,

$$E[Y \mid \Theta = \theta_k^{(n)}] = \text{midpoint of } A_k^{(n)}.$$

Here $E[X \mid A]$ denotes the expectation of a random variable X, given that event A has occurred. (4.10) requires that partitioning intervals corresponding to examinees expected to get "approximately" half right on any item of the assessment test should be removed. Since the M items are fixed and ε in (4.10) can be as small as desired, (4.10) will always be true in practice. But nonetheless some Monte Carlo analysis here should help support the author's intuition that the practical impact of ignoring (4.10) in practice should be minimal. Alternatively one could merely throw out those intervals $A_k^{(n)}$ producing $\hat{p}_i^{(k)}$'s exceedingly close to $\frac{1}{2}$ for any $i = n+1,\dots,n+M$. The mathematically interested reader is invited to consult p. 119 Serfling (1980) for the need of an assumption such as (4.10).

This completes the list and discussion of the assumptions required to obtain the large sample theory when $d = 1$. In summary, all the assumptions either seem quite reasonable or at least, in some cases, not unduly restrictive. Of course, Monte Carlo studies and the use of the procedure in actual mental test data are needed to verify this opinion (see Section 7).

Now the two basic large sample results when $d = 1$ are stated.

<u>Theorem 4.1</u>. Suppose (4.1), (4.2), M fixed assessment items, continuous and differentiable ICC's, (4.5), (4.6), (4.7), all included subintervals in $(\eta,1]$, (4.9), and (4.10). Then (see (3.7)), for any choice of integers $k \equiv k_n$

$$(4.11) \qquad Z_n \equiv \frac{(\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k})\sqrt{J_k^{(n)}}}{\left[(\hat{\mu}_4(Y) - \hat{\sigma}^4_{Y,k}) - \dfrac{\hat{\delta}_{4,k}}{M^4}\right]^{\frac{1}{2}}} \xrightarrow{L} Z \quad \text{as} \quad n \to \infty,$$

where the rate of convergence of the distribution function of Z_n to a $N(0,1)$ distribution function does not depend on the sequence $\{k_n\}$.

Theorem 4.2. Suppose (4.1), (4.2), M fixed assessment items, continuous and differentiable ICC's, (4.5), (4.6), (4.7), all included subintervals in $(\eta,1]$, (4.8), and (4.10). Then (see 3.7),

$$(4.12) \qquad T_n \xrightarrow{L} Z \quad \text{as} \quad n \to \infty.$$

Remark. The denominator of (4.11) has a rather tidy matrix representation: Let, suppressing k, and writing $\hat{p}_i$ for $\hat{p}_{i+n}^{(k)}$ of (3.4) for convenience,

$$\hat{D} \equiv (2\hat{p}_1-1, 2\hat{p}_2-1, \ldots, 2\hat{p}_M-1, 1)$$

and

$$\hat{\Sigma} = \begin{vmatrix} \hat{p}_1(1-\hat{p}_1) & 0 & \cdots & 0 & \hat{p}_1(1-\hat{p}_1)(1-2\hat{p}_1) \\ 0 & \hat{p}_2(1-\hat{p}_2) & \cdots & 0 & \hat{p}_2(1-\hat{p}_2)(1-2\hat{p}_2) \\ \vdots & 0 & & \vdots & \vdots \\ & \vdots & & 0 & \\ 0 & 0 & \cdots & \hat{p}_M(1-\hat{p}_M) & \hat{p}_M(1-\hat{p}_M)(1-2\hat{p}_M) \\ \hat{p}_1(1-\hat{p}_1)(1-2\hat{p}_1) & \hat{p}_2(1-\hat{p}_2)(1-2\hat{p}_2) & \cdots & \hat{p}_M(1-\hat{p}_M)(1-2\hat{p}_M) & M^4(\hat{\mu}_4(Y) - \hat{\sigma}_Y^4) \end{vmatrix}$$

Then, it is trivial to see that

$$(4.13) \quad (M^4(\hat{\mu}_4(Y) - \hat{\sigma}^4_{Y,k}) - \hat{\delta}_{4,k})^{\frac{1}{2}} = (\hat{D}\hat{\Sigma}\hat{D}')^{\frac{1}{2}}.$$

This may be useful both for interpretation and computer programming purposes.

<u>Proofs</u>. The proofs of these two results are lengthy, delicate, and make use of sophisticated results from mathematical probability theory. Their presentation here seems inappropriate. The interested reader is referred to Appendix I for a complete presentation of the proofs. However, a brief sketch of the highlights of the proofs seems appropriate. We will consider Theorem 4.1's proof only, since that of Theorem 4.2 is rather similar.

Basically, the proof is based upon the central limit theorem holding for sums of independent finite variance random variables, the summands suitably similar in magnitude (see e.g., the Lindeberg-Feller theorem as stated for triangular arrays of random variables in Chung (1974), p. 205). It is also based on a multivariate version of the "δ method", the local linearization technique based on multivariate Taylor series expansion that is so useful in large sample theory (see Serfling (1980), p. 122). Further, the random denominator in (4.11) is permitted by an application of a Slutsky type argument (see Serfling (1980), p. 19). Finally, a very delicate question about the asymptotic expectation of $\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{P,k}$ is dealt with by means of a conditional tail probability estimate for the random ability Θ, given that Y (the partitioning score) $\in A_k^{(n)}$. This conditional probability estimate follows in a manner similar to the classical exponential probability bounds for sums of independent random variables (see, e.g., Stout (1974), Section 5.2) and is the most delicate aspect of the proof.

5. The concept of latent trait dimensionality

The purpose of this section is to propose a concept of dimensionality distinct from the classical item response theory definition. Recall that the classical definition amounts to stating that k is the correct dimensionality provided the joint distribution of the observed random variables $\{U_{ij}, 1 \leq i \leq N, 1 \leq j \leq J\}$ is representable in terms of a k dimensional distribution of iid $\underline{\Theta}_j$ and identical (in j) conditional distributions of $\underline{U}_j$, given $\underline{\Theta}_j$, with local independence and (3.4) assumed, and moreover that such a representation is impossible for $k' < k$.

Although mathematically appealing, this definition is rather inappropriate for assessing the power of tests of dimensionality in the case of mental test theory. This is because in mental testing, individual test items clearly have multiple determinants of their respective probabilities of correct response, thus necessitating that $\underline{\theta}$ in the $P_i(\underline{\theta})$'s be multidimensional. This position is perhaps most clearly and vigorously pursued by Humphreys (see Humphreys, 1984). Humphreys states:

> "The related problems of dimensionality and bias of items are being approached in an arbitrary and over-simplified fashion. It should be obvious that unidimensionality can only be approximated. Even in highly homogeneous tests the mean correlation between paired items is quite small. The large amount of unique variance in items is not random error, although it can be called error from the point of view of the attribute that one is attempting to measure. Test theory must cope with these small correlations. We start with the assumption that responses to items have many causes or determinants."

Humphreys (1984) presents the viewpoint that dominant attributes (dimensions) result from overlapping attributes common to many items. Attributes unique to individual items or common to relatively few items are unavoidable and indeed are not detrimental to the measurement of dominant dimensions. But the number of these attributes should not be counted in assessing the "dimensionality" of the test. Humphreys' writings stress that the observed low item intercorrelations compel one to accept the viewpoint of multiply determined items. Unfortunately, the classical definition takes as test dimensionality the total number k of all item dimensions, each item requiring in general several (possibly many) dimensions to describe its $P_i(\underline{\theta})$. As follows from the above discussion of Humphreys' viewpoint, this fact is true <u>even</u> in situations where both from a psychometric and a data analytic viewpoint, one would want to catagorize tests as unidimensional. Thus the classical definition assigns dimensionality $d = k > 1$ (k thus assigned possi quite large in fact) in settings where one would want to assign $d = 1$. The following hypothetical example is intended to make concrete the multi-dimensional nature of items in tests that should be considered unidimensional.

<u>Example 5.1.</u> Consider a "probability" test where Item 1 measures ability in probability but is influenced by many other idiosyncratic factors contributing to "non-error noise", e.g., a knowledge of the rules of bridge.

Item 2 measures ability in probability but is influenced among other things by the examinees' understanding of elementary physics.

Item 3 measures ability in probability but is influenced among other things by a knowledge of Mendelian genetics.

.
.
.

One clearly is forced to label such a test as multidimensional according to the classical psychometric conceptualization of dimensionality described above (indeed, clearly $d \geq 4$ will be assigned with the dimensions including ability in probability, bridge knowledge, elementary physics knowledge, and knowledge of genetics).

Clearly one wants a conceptualization of test dimensionality d such that tests such as that of Example 3.1 would be considered one dimensional (i.e., $d = 1$). What is needed is a conceptualization of dimensionality that does not yield an inflated d as a result of the inherent multidimensionality of items.

To this end, the following conceptualization of dimensionality is proposed. Consider a test of length $nd + M$ administered to one randomly selected examinee yielding random variables $\{U_i, 1 \leq i \leq nd + M\}$, where n is to be thought of as possibly large compared to M. Let $\underline{\Theta}$ have a k dimensional density. (Here k will likely in applications be quite large because of multiply determined items.) Let

$$(5.1) \qquad Y_1 = \frac{\Sigma_{i=1}^{n} U_i}{n}, \quad Y_2 = \frac{\Sigma_{i=n+1}^{2n} U_i}{M}, \ldots, Y_d = \frac{\Sigma_{i=n(d-1)+1}^{nd} U_i}{M},$$

$$Y_{d+1} = \frac{\Sigma_{i=nd+1}^{nd+M} U_i}{M}$$

define a splitting of the test into d "partitioning" subtest scores and "assessment" subtest score Y_{d+1}. Let, for $0 \leq y \leq 1$,

$$(5.2) \quad A_{\underline{y}} \equiv A_{y_1,\ldots,y_d} = \{\underline{\theta} : E[Y_1 | \underline{\Theta} = \underline{\theta}] = y_1, E[Y_2 | \underline{\Theta} = \underline{\theta}] = y_2, \ldots, E[Y_d | \underline{\Theta} = \underline{\theta}] = y_d\}.$$

Let

$$(5.3) \quad s = \underbrace{\int_0^1 \cdots \int_0^1}_{d \text{ integrals}} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{k \text{ integrals}} (E[Y_2 \mid \underline{\Theta} = \underline{\theta}] - E[Y_2 | \underline{\Theta} \in A_{\underline{y}}])^2 f(\underline{\theta} | \underline{\Theta} \in A_{\underline{y}}) d\underline{\theta} d\underline{y},$$

where $f(\underline{\theta} \mid \underline{\Theta} \in A_{\underline{Y}})$ denotes the conditional density of $\underline{\Theta}$, given $\underline{\Theta} \in A_{\underline{Y}}$. Relative to the particular splitting that produced $Y_1, Y_2, \ldots, Y_{d+1}$, s should be viewed as an index measuring the lack of fit of the assumption that the dimensionality is d. Let S denote the set of the

$$\binom{nd+M}{\underbrace{n\ n\ \ldots\ n}_{d\ n\text{'s}}\ M}$$

possible such splittings. Let

$$S_d = \sup_{s \in S} s \,. \tag{5.4}$$

Definition 5.1. A test is said to be of dimension d provided $S_d = 0$, $S_{d-1} > 0$.

In actual applications, taking into account the inherent multidimensionality of individual items, it seems reasonable, indeed necessary to assign dimensionality d provided S_d is "small" (but not necessarily 0) and S_d is "not small", the quantification of "small" and "not small" varying from application to application. When this last remark is taken into account, Definition 5.1 is much in the spirit of the factor analytic conceptualization of dimensionality and is precisely in the spirit of the Humphreys' viewpoint of dimensionality. That is, the dimensionality is taken to be the number of common factors (dominant attributes) with specific factors (attributes common to relatively few items) not contributing to the dimensionality. The simple hypothetical example below is intended to illustrate this admittedly mathematically complex but actually rather intuitive definition.

<u>Example 5.2</u>. Definition 5.1 will be applied with $d = 1$. That is, the question whether $d = 1$ in the sense of Definition 5.1 will be addressed.

<u>Case</u> 1. Suppose that two types of examinees, "rural" and "urban", take a "reading test". Let the classical dimensionality $k = 2$ with θ_1 denoting the level of reading ability and θ_2 denoting familiarity with urban culture. Suppose, for simplicity, that

$$P[\underline{\Theta} = (-1,1)] = P[\underline{\Theta} = (1,-1)] = \tfrac{1}{2},$$

defines the latent ability space. Here

$(\theta_1,\theta_2) = (-1,1)$ denotes low reading ability, high familiarity with urban culture

and

$(\theta_1,\theta_2) = (1,-1)$ denotes high reading ability, low familiarity with urban culture.

Consider the computation of S_d when $d = 1$ in Definition 5.1. I.e., consider the question of unidimensionality in the sense of Definition 5.1. Suppose

$$E[Y_1 \mid \underline{\Theta} = (-1,1)] = E[Y_1 \mid \underline{\Theta} = (1,-1)] = \tfrac{1}{2}$$

for a particular splitting to obtain subtest scores Y_1, Y_2. That is, both types of examinees ((-1,1) and (1,-1)) can be expected to on the average get $\tfrac{1}{2}$ of the items on the partitioning subtest right. Thus

$$A_{\frac{1}{2}} = \{(-1,1),(1,-1)\}.$$

Suppose that the assessment score Y_2 has been formed from M items with identical item characteristic curves, each item dependent on θ_1 only (i.e. the items upon which Y_2 is based really are pure "reading" items.

In particular, suppose

$$P_i(-1) = 1/4, \qquad P_i(1) = 3/4$$

for $i = n+1,\ldots,n+M$ defines $P_i(\Theta_1)$. (The range of Θ_1 is $\{-1,1\}$). Since this reading test is clearly contaminated by a second dimension, $d = 2$ should clearly be concluded. Replacing the integrals by sums in (5.4) because of the artificial discrete nature of this model, one obtains

(5.5) $S_1 \geq s = (E[Y_2 \mid \underline{\Theta} = (-1,1)] - \tfrac{1}{2})^2 \tfrac{1}{2} + (E[Y_2 \mid \underline{\Theta} = (1,-1)] - \tfrac{1}{2})^2 \tfrac{1}{2} = 1/16.$

Thus $S_1 > 0$, reflecting the fact that $d > 1$ (in fact $d = 2$) in this case.

Case 2. Now by contrast, consider a case where reading ability is uncontaminated by a second dimension. Suppose in fact that all $n + M$ items satisfy for $1 \leq i \leq n + M$,

$$P_i(-1) = \tfrac{1}{4}, \quad P_i(0) = \tfrac{1}{2}, \quad P_i(1) = 1$$

where the latent ability space (reading ability here denoted by Θ_1) is $\{-1,0,1\}$. Here $k = 1$. Clearly $A_{1/4} = \{-1\}$, $A_{1/2} = \{0\}$, and $A_1 = \{1\}$ and

$$S = \sup_{S} s = 0,$$

leading one to conclude that $d = 1$, as is desired in this situation.

The above example did not have the classical dimensionality k large as the result of multiply determined items. However, it is intuitively clear that for each test splitting, s as defined by (5.3) should be small in Example 5.1 since there is only one dominant attribute there. For, conditioning on $\underline{\Theta} \in A_y$ intuitively amounts to conditioning on a subpopulation of examinees of roughly the same probability ability (indexed by θ_1 say) but with possibly widely differing abilities on the other $k - 1$ components of $\underline{\Theta}$. Thus for "typical" $\underline{\theta}$'s, given $\underline{\Theta} \in A_y$, $(E[Y_2|\underline{\Theta}=\underline{\theta}] - E[Y_2 \mid \underline{\Theta} \in A_y])^2$ should be small since $E[Y_2|\underline{\Theta}=\underline{\theta}]$ should be mainly influenced by θ_1 with the effect of $\theta_2,\ldots,\theta_k$ on $E[Y_2|\underline{\Theta}=\underline{\theta}]$ being small. Hence, as suggested, S_1 should be small for situations like Example 5.1. One can construct examples showing that this is indeed the case. The following example is such an example. That is, the classical dimensionality greatly exceeds 1 because of multiple determinants and yet the dimensionality according to Definition 5.1 is 1. This example is highly artifical and is merely intended to make concrete the above comments.

Example 5.3. Suppose the N ICC's are given by

$$P_i(\underline{\theta}) = \theta_1 + \theta_i \quad \text{where } \frac{1}{4} \le \theta_1 \le \frac{3}{4}$$

and $-\frac{1}{4} < \theta_i < \frac{1}{4}$ for $i = 2,\ldots,N$. Suppose each Θ_i is uniformly distributed over its range. Fix M. Split the test into subtests of size M, $N - M = n$. Note that

$$E[Y_2|\underline{\Theta} = \underline{\theta}] = \theta_1, \quad E[Y_1|\underline{\Theta} = \underline{\theta}] = \theta_1$$

Hence

$$E[Y_2|\underline{\Theta} \in A_y] = E[Y_2|\Theta_1 = \theta_1 \text{ and } (\Theta_2,\ldots,\Theta_{N+1}) \text{ unrestricted}]$$

$$= \theta_1.$$

Thus recalling (5.2), (5.3), (5.4), $S_1 = s = 0$. Hence $d = 1$, even though $k = N + 1$.

Some final remarks in aiding the reader in interpreting (5.2) - (5.4) seem in order. One should consider $d = 1$ for ease of understanding and k small as in Example 5.2. Moreover, one should realize that (taking $d = 1$) the integration over y is merely an averaging process over y of the contribution to multidimensionality resulting from each choice of y such that $E[Y_1 \mid \underline{\Theta} = \underline{\theta}] = y$.

Put qualitatively, s measures, among examinees that are expected to have the same score on the partitioning subtest, the amount of variation there is in their expected scores on the assessment subtest.

6. <u>A discussion of power</u>. The preceeding section suggests that if S_1 (see (5.4)) is sufficiently large, then one would like to reject the hypothesis of unidimensionality with high probability. The following result establishes ·ch a property for long tests and large examinee samples.

<u>Theorem 6.1</u>. Let $d = 1$ in (5.3). Suppose for latent ability $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ that

(6.1) the density $f(\underline{\theta})$ is a continuous function of $\underline{\theta}$,

(6.2) each ICC $P_i(\underline{\theta})$ is continuous in $\underline{\theta}$,

(6.3) there exists $\varepsilon > 0$, $M \geq 1$, and some test splitting into subsets for each n such that

$$s > \frac{1 + \varepsilon}{M} \quad \text{for each} \quad n \geq 1.$$

Then, for fixed level of significance α, and prescribed power β, $0 < \beta < 1$,

(6.4) $$P[\text{reject } H] = P[T > Z_\alpha] > \beta$$

for n sufficiently large.

<u>Proof</u>. See Appendix II for the proof.

The following corollary is useful in interpreting psychometrically the meaning of (6.1). Let $Y_j \equiv Y_{j,p}^{(k)}$ be the partitioning subtest score of the jth examinee of the partitioning interval $A_k^{(n)}$.

<u>Corollary 6.1</u>. Assume all the hypotheses of Theorem 6.1, except for (6.8). Suppose for a particular test splitting that for all n and some $\varepsilon > 0$

$$(6.5)\qquad \frac{\sum_{k=1}^{K_n} \sum_{i=1}^{M} \sum_{i'=i+1}^{M} \operatorname{cov}(U_{i1k}, U_{i'1k})}{K_n \binom{M}{2}} \geq \varepsilon$$

Then (6.3) holds and hence (6.4) holds for n sufficiently large.

Proof. See Appendix II for the proof.

The interpretation of the corollary is that if the average of the item intercovariances on the assessment subtest is bounded away from 0, then the power can be made arbitrarily close to 1 for large n. Here the averaging is over both subintervals of the partition and over item pairs of the assessment test. The idea is that if conditioning still leaves a significant amount of interitem covariance on the average, then one would want to reject with reasonable power. Corollary 6.1 guarantees this.

7. Preliminary Monte Carlo studies: As of the writing of this report, an extensive Monte Carlo study has been begun. Its purpose is to assess the performance of the proposed statistical test of unidimensionality for values of test length, examinee sample size and subtest splitting sizes typically encountered in actual test settings where dimensionality is an issue. A part of this study will consist of the consideration of minor modifications of the proposed procedure in order to improve its performance for small test length N (and, as turns out to be potentially useful, "large" M/N).

Preliminary Monte Carlo studies clearly suggest the feasibility of the test procedure and/or minor modifications of it for actual mental test applications. The purpose of this section is to summarize these preliminary findings. A Fortran program to do Monte Carlo simulation of the performance of the test procedure (given by (3.8)) both when $d = 1$ and when $d = 2$ has been written. Further, a method of introducing multiply determined items has been programmed, thus handling the $d = 2$ case when the classical dimensionality $k > 2$, as well as handling the $k = d = 2$ case.

Briefly, in the $d = 1$ case, the standard three parametric logistic model is used with $c_i = 1/5$ for all i and parameters (a_i, b_i) randomly generated to simulate values typically occurring in applications. Two results obtained were that the procedure showed good adherence to the prescribed level of significance when

(7.1) $J = 1000, \quad N = 200, \quad M = 10$

and when

(7.2) $J = 1000, \quad N = 75, \quad M = 5.$

Using a bivariate logistic model for the $d = 2$ case, the procedure displayed excellent power for the (7.1) and (7.2) cases.

The introduction of multiply determined items produced only insignificant loss of power and a modest increase in the actual level of significance α for both the (7.1) and (7.2) cases.

Further, as a result of the preliminary Monte Carlo studies, there is evidence that a minor modification of the denominator of (3.7), namely replacing

$$\left\{\sum_{k=1}^{K_n}\left[(\hat{\mu}_4(Y) - \hat{\sigma}^4_{Y,k}) - \frac{\hat{\delta}_{4,k}}{M^4}\right]/J_k^{(n)}\right\}^{1/2}$$

by

$$\left\{\sum_{k=1}^{K_n}\left[(\hat{\mu}_4(Y) - \hat{\sigma}^4_{Y,k}) + \frac{\hat{\delta}_{4,k}}{M^4} + 2\sqrt{(\hat{\mu}_4(Y) - \hat{\sigma}^4_{Y,k})\hat{\delta}_{4,k}/M^4}\right]/J_k^{(n)}\right\}^{1/2},$$

may considerably improve the performance of the procedure for small test length N, that is, produce good adherance to the prescribed level of significance with good power performance still preserved. Use of this denominator is suggested by the well known Cauchy-Schwarz inequality from mathematics. This and other modifications of the proposed procedure are presently being considered.

As the theory suggests, preliminary Monte Carlo studies indicate the accuracy of the large sample approximation begins to break down when M/N becomes too large. E.g., $N = 100$, $M = 10$, $J = 1000$ seemed to noticeably inflate the actual level of significance. Various possible modifications of the procedure may largely eliminate this breakdown.

The Fortran coding, together with instructions, is obtainable from the author. Again, the preliminary nature of these Monte Carlo findings is to be emphasized.

8. <u>The choice of assessment test length and assessment test items</u>: Consider a mental test setting consisting of an N item test. The choice of M should be "small" compared to test length N; the preliminary Monte Carlo studies reported on in Section 7 tentatively suggest taking $M/N < 1/10$. On the other hand, it is also clear that in a reasonably long test (e.g., $N \geq 100$), the larger that one can safely take M, and thus the greater statistical power one attains. In a particular test application with known N and J, it would be feasible for an applicator to conduct a Monte Carlo simulation for various choices of M in order to select an M that is large enough to provide good power, yet not so large that the actual level of significance deviates too much from the prescribed level α.

An even more critical issue is the assignment of test items to the assessment subtest, once M has been determined. If done improperly, the statistical test may have little or no statistical power. The problem is essentially that, even though it may be true that $d > 1$ in the sense of Definition 5.1, nonetheless, the assignment of assessment items may be such that the two subtests are "parallel" in a certain sense. That is, letting Y_1 and Y_2 denote the proportions correct on the partitioning and assessment subtests respectively,

$$(8.1) \qquad EY_2 = h(EY_1)$$

may be the case for each examinee, when h is a strictly increasing function that is the same for each examinee. Then in fact the statistical test will have no power. Indeed, if the two subtests are roughly parallel in the sense that (8.1) holds approximately then there will be little or no statistical power. Hence, in order that the statistical test have reasonable power, the items must be assigned to the two subtests in such a way that if indeed

$d > 1$ is the case then there should be little probability that the two subtests will be approximately parallel.

The procedure for assigning items to the assessment subtest should be such that, when there are other dominant dimensions present, then the items that are more heavily influenced by these other dimensions, should (ideally) all be assigned to one subtest. And hence the items assigned to the other subtest should be less heavily influenced by these other dimensions.

For the assignment of items to the assessment subtest, three methods that could be used are:

(a) Expert judgment: One or more experts on mental testing could select a set of M items that seem to display a common bias or seem to be significantly influenced by a second dimension. This choice could be made on the basis of item content, method of item administration, item format, past experience from previous test administration, subjective impressions, etc.

(b) Factor analytic or other multivariate analyses: Split the text examinee population (randomly) into two subpopulations. Use some multivariate technique on the first examinee subpopulation to select those items that seem to most represent other unified dimensions than the dominant test dimension. For example, using the sample tetrachoric correlation matrix, one could perform the usual factor analysis of the test. Then those M items with the largest estimated positive loadings (one could use largest negative loadings just as well) on other dimensions of the factor analysis relative to the dominant dimension should be assigned to the assessment subtest. Then carry out the statistical test of Section 3 on the second examinee subpopulation.

The point to realize here is that even though there is not an adequate theoretical or empirical basis to justify the use of factor analysis to assess the dimensionality of dichotomously scored items; nonetheless, factor analysis can be used as an atheoretical data-analytic tool in conjunction with the theoretically grounded procedure presented in this paper. Since often test population size is quite large in test applications, there seems to be much to recommend this approach.

(c) Random assignment of items to subtests, subject to user specification of the magnitude of M,n. The hope here is that the diversity of item dependence on the various dimensions in the $d > 1$ case would with high probability bring about the selection of sufficiently non-parallel subtests that the statistical test will display reasonable power. At present, no theoretical nor Monte Carlo work has been carried out to defend this hope. Indeed, the author is skeptical of the effectiveness of this method of assigning items to subtests, but will, of course investigate it.

One has to choose the partition subintervals. These intervals should clearly be of equal width. Hence, the user is really only selecting the number of subintervals K_n. Roughly, these intervals should be as narrow as possible subject to the requirements that a reasonable number of intervals (e.g., at least 10) have a reasonable number of examinees (say at least 25) assigned to them and that not too many examinee scores are discarded. One final minor matter is that a convention for excluding intervals with too examinees assigned to them must be selected. E.g., one might require at least 25 examinees/interval.

After the completion of the extensive Monte Carlo study referred to at the beginning of Section 7, the author should be able to provide clearer guidelines on the user specified options discussed here in Section 8 that are required to conduct the statistical test.

9. Extension to tests of higher dimensionality. The procedure can be modified to test

$$H: d \leq d_0 \text{ vs } A: d > d_0$$

for any $d_0 \geq 2$. One merely uses d_0 partitioning subtests and one assessment subtest and then partitions examinees on the basis of what partitionery subset of the d_0 dimensional unit cube their partitioning score $Y_1, Y_2, \ldots, Y_{d_0}$ falls into. The author expects to do some theoretical and Monte Carlo work on the $d_0 \geq 2$ case in the future.

10. Concluding remarks. It should be stressed that the main advantage of the proposed procedure of this paper is that it has a rigorous asymptotic theory backing it up. As a result, it should display good power against those alternatives for which one really does wish to reject the hypothesis of unidimensionality and it should not spuriously reject unidimensionality in situations where one does not wich to reject, as has sometimes been the case with other procedures. Further, the assumptions required for the asymptotic theory are rather unrestrictive and are totally non-parametric. That is, no particular form, such as multivariate logistic or multivariate normal is assumed for the item response theory model.

Appendix I. Proofs of Theorems 4.1 and 4.2.

Let $s_k^{(n)}$ denote the radius of $A_k^{(n)}$. Let $\varphi_k^{(n)}$ denote the midpoint of $A_k^{(n)}$. Let $\theta_k^{(n)}$ be defined by, writing $\theta_k^{(n)}$ instead of $\Theta_1 = \theta_k^{(n)}$ for the conditioning event,

$$(I.1) \qquad E[Y_1^{(k)} | \theta_k^{(n)}] = \varphi_k^{(n)},$$

noting that $\theta_k^{(n)}$ is well defined for all $1 \leq k \leq K_n$, $n \geq 1$ by the continuity and strict increase of the $P_i(\theta)$ and the definition of η in Section 4. Suppose, for the ε of (4.5) and (4.7), that

(I.2) $\theta_k^{(n)} \in (a + \varepsilon, b - \varepsilon)$ for all k,n .

Recall that the density of Θ is denoted by $f(\theta)$. The conditional probability estimate given in Proposition 4.1 below is central to the derivati of the asymptotic theory. It is based on the Berry-Esseen theorem, which is now stated for completeness. Let $V(W)$ denote the variance of a random variable and $\Phi(z)$ denote the distribution function of a $N(0,1)$ random variable throughout.

Lemma I.1. Fix n. If $\{W_i, 1 \le i \le n\}$ are independent mean 0 random variables with

$$\Sigma_{i=1}^{n} V(W_i) > 0$$

then, letting $S_n = \Sigma_{i=1}^{n} W_i$, $\Gamma_n = \Sigma_{i=1}^{n} E|W_i|^3$, there exists a universal constant C such that

$$\sup_x |P[\frac{S_n}{V(S_n)^{1/2}} < x] - \Phi(x)| \le \frac{C\Gamma_n^3}{V(S_n)^{3/2}}$$

Proposition I.1. Fix k,n and denote $s_n \equiv s_k^{(n)}, \theta^{(n)} \equiv \theta_k^{(n)}$. Suppose

(I.3) $2n^{-1} \le s_n \le C\, n^{-1/2}$.

Then, letting $Y \equiv Y_p$ denote the proportion correct on the partitioning subtest for a randomly chosen examinee,

(I.4) $$P[|\Theta - \theta^{(n)}| \ge \frac{2x}{\varepsilon} \mid Y \in A_k^{(n)}] \le \frac{C\exp(-n\, x^2/4)}{s_n(n^{1/2}s_n - n^{-1/2})}$$

for all $s_n \le x \le (\log n)^{-1/2}$ and ε of (4.5) and (I.2).

Proof. Clearly

(I.5) $$P[|\Theta - \theta^{(n)}| \ge \frac{2x}{\varepsilon}| Y \quad A_k^{(n)}] \doteq \frac{P[|\Theta - \theta^{(n)}| \ge \frac{2x}{\varepsilon}, |Y - \varphi^{(n)}| \le s_n]}{P[|Y - \varphi^{(n)}| \le s_n]} .$$

The denominator is estimated first. For any $y > 0$,

(I.6) $$P[|Y - \varphi^{(n)}| \le s_n] \ge P[|Y - \varphi^{(n)}| \le s_n, |\Theta - \theta^{(n)}| \le y]$$
$$= \int_{\theta^{(n)}-y}^{\theta^{(n)}+y} P[|Y - \varphi^{(n)}| \le s_n|\Theta = \theta]f(\theta)d\theta$$
$$\ge \int_{\theta^{(n)}-y}^{\theta^{(n)}+y} P[|Y - E[Y|\theta]| + |\varphi^{(n)} - E[Y|\theta]| \le s_n|\Theta = \theta]f(\theta)d\theta \equiv \gamma$$

say. Now for $\theta \in [\theta^{(n)} - y, \theta^{(n)} + y]$,

$$|\varphi^{(n)} - E[Y|\theta]| = |E[Y|\theta^{(n)}] - E[Y|\theta]|$$
$$= \left|\frac{\Sigma_{i=1}^{n} \{P_i(\theta^{(n)}) - P_i(\theta)\}}{n}\right| \le C|\theta^{(n)} - \theta|$$

by the mean value theorem and (4.6) Thus, continuing (I.6),

(I.7) $$\gamma \ge \int_{\theta^{(n)}-y}^{\theta^{(n)}+y} P[|Y - E[Y|\theta]| \le s_n - Cy|\theta]f(\theta)d\theta$$
$$\ge \int_{\theta^{(n)}-y}^{\theta^{(n)}+y} P[|Y - E[Y|\theta]| \le s_n/2|\theta]f(\theta)d\theta ,$$

choosing

(I.8) $$y \le \frac{s_n}{2C} .$$

Now, trivially, writing $U_i \equiv U_{ijk}$

(I.9) $$V(Y|\theta) \le n^{-1} , \Sigma_{i=1}^{n} E|U_i - E[U_i|\theta]|^3 \le n .$$

Further, by (4.7), (I.8), (I.2), (I.3) it follows that

(I.10) $$V(nY|\theta) = \Sigma_{i=1}^{n} P_i(\theta)(1 - P_i(\theta)) \ge \varepsilon n \text{ for some } \varepsilon > 0.$$

Hence, by the Berry-Esseen Theorem (Lemma I.1), noting (I.9), (I.10) and

(I.3), it follows that

$$P[|Y - E(Y|\theta)| \leq \frac{s_n}{2} |\theta] \geq P[\frac{|Y - E(Y|\theta)|}{\sqrt{V(Y|\theta)}} \leq \frac{n^{1/2}s_n}{2} |\theta]$$

$$\geq \frac{-C}{n^{1/2}} + \Phi\left(\frac{s_n n^{1/2}}{2}\right) - \Phi\left(\frac{-s_n n^{1/2}}{2}\right) \geq -\frac{C}{n^{1/2}} + Cs_n n^{1/2}$$

Hence, using (I.6) and (I.7),

$$\text{(I.11)} \qquad P[|Y - \varphi^{(n)}| \leq s_n] \geq C s_n (s_n n^{1/2} - n^{-1/2}) .$$

Now, let $x' = 2x/\varepsilon$ and consider the numerator of the right hand side of (I.5) :

$$\text{(I.12)} \qquad P[|Y - \varphi^{(n)}| \leq s_n, |\Theta - \theta^{(n)}| \geq x']$$

$$= \int_{|\theta-\theta^{(n)}| \geq x'} P[|Y - \varphi^{(n)}| \leq s_n |\Theta = \theta] f(\theta) d\theta .$$

Assume

$$\text{(I.13)} \qquad |\theta - \theta^{(n)}| \geq x', a < \theta < b .$$

Now

$$\text{(I.14)} \qquad \delta \equiv P[|Y - E[Y|\theta^{(n)}]| \leq s_n |\theta] \leq P[|Y - E[Y|\theta^{(n)}]| \leq x|\theta]$$

$$\leq P[|Y - E[Y|\theta]| \geq |E[Y|\theta] - E[Y|\theta^{(n)}]| -x$$

since, on $[|Y - E[Y|\theta^{(n)}]| \leq x]$,

$$x \geq |Y - E[Y|\theta^{(n)}]| \geq |E[Y|\theta] - E[Y|\theta^{(n)}]| - |Y - E[Y|\theta]| .$$

Now, using (4.5) and (I.2), it follows that

$$|E[Y|\theta] - E[Y|\theta^{(n)}]| \geq \varepsilon |\theta - \theta^{(n)}| \geq \varepsilon x' = 2x$$

Hence, using (I.12) and (I.14)

$$\text{(I.15)} \qquad \delta \le P[|Y - E[Y|\theta]| \ge x|\theta] = P\left[\frac{|\Sigma_1^n U_i - E[U_i|\theta]|}{V(\Sigma_{i=1}^n U_i|\theta)^{1/2}} \ge \frac{nx}{V(\Sigma_{i=1}^n U_i|\theta)^{1/2}}\middle|\theta\right]$$

Now,

$$\frac{|U_i - E[U_i|\theta]|}{V(nY|\theta)^{1/2}} \le \frac{1}{V(\Sigma_{i=1}^n U_i|\theta)^{1/2}} \equiv c'.$$

Let

$$\varepsilon' = \frac{nx}{V(\Sigma_{i=1}^n U_i|\theta)^{1/2}} .$$

Note that

$$\text{(I.16)} \qquad \varepsilon' c' = \frac{nx}{V(\Sigma_{i=1}^n U_i|\theta)} \to 0 \quad \text{as} \quad n \to \infty$$

by (4.7) and the hypothesis that $x \le (\log n)^{-1/2}$. But, this means that conditional on $\Theta = \theta$, the U_i satisfy the conditions of the classical Kolmogorov exponential bounds (see, e.g., Stout (1974), p. 262). That is,

$$\delta \le \exp\left[-\frac{(\varepsilon')^2}{2}(1 - \varepsilon' c')\right]$$

from which it follows by (I.16) and (I.9) for n large that for θ such that $|\theta - \theta^{(n)}| \ge x', \theta \in (a,b)$,

$$\text{(I.17)} \qquad P[|Y - E[Y|\theta^{(n)}]| \le s_n|\Theta = \theta] \le \exp\left(\frac{-x^2 n}{4}\right) .$$

Now assume instead that θ satisfies

$$\text{(I.18)} \qquad |\theta - \theta^{(n)}| \ge x', \ \theta \notin (a,b).$$

Then, for $\theta > b$, using the mean value theorem and (4.5),

$$E[Y|\theta] - E[Y|\theta^{(n)}] \geq E[Y|b] - E[Y|\theta^{(n)}] \geq \varepsilon^2 .$$

Thus, using a similar argument for $\theta < a$,

$$E[Y|\theta] - E[Y|\theta^{(n)}]| \geq \varepsilon^2 .$$

Hence, reasoning as in (I.14), it follows, for n large that

$$\delta \leq P[|Y - E[Y|\theta]| \geq \varepsilon^2/2 | \Theta = \theta] .$$

Hence, for any $\varepsilon_n \to 0$, for n large

$$\text{(I.19)} \qquad \delta \leq P[|Y - E[Y|\theta]| \geq \varepsilon_n | \Theta = \theta].$$

Now, it is straightforward to verify (see Lamperti (1966), p. 44)) for $0 \leq t \leq 2$ that

$$\text{(I.20)} \qquad E[\exp(t\, n(Y - E[Y|\theta])|\theta] \leq \exp\left[\frac{\Sigma_{i=1}^{n} P_i(\theta)(1 - P_i(\theta))}{2} t^2(1 + t)\right].$$

Then, by Markov's inequality, for $0 \leq t \leq 2$,

$$\text{(I.21)} \quad P[n|Y - E[Y|\theta]| \geq n\,\varepsilon_n|\theta] \leq 2\exp(-n\,\varepsilon_n\, t)\exp\left[\frac{\Sigma_{i=1}^{n} P_i(\theta)(1 - P_i(\theta))}{2} t^2($$

There are two cases to consider. First, suppose that n is such that

$$\text{(I.22)} \qquad \Sigma_{i=1}^{n} P_i(\theta)(1 - P_i(\theta)) \geq 2n/\log n .$$

Then (see (7), p. 44 of Lamperti (1966)) choosing

$$t = \frac{2\,\varepsilon\, n}{\Sigma_{i=1}^{n} P_i(\theta)(1 - P_i(\theta))} , \quad \varepsilon_n = \frac{1}{\log n}$$

yields

$$\text{(I.23)} \qquad \delta \leq P[n|Y - E[Y|\theta]| \geq \frac{n}{\log n}|\ \theta\ ] \leq 2\ \exp[\frac{n}{16\ \log n}] .$$

Suppose, instead of (I.22) that

$$\sum_{i=1}^{n} P_i(\theta)(1 - P_i(\theta)) < \frac{2n}{\log n} \ .$$

Then, letting $t = \frac{1}{8}$ and using (I.21) yields

$$\text{(I.24)} \quad \delta \le P[n|Y - E[Y|\theta]| \ge \frac{n}{\log n} |\theta] \le 2\exp(\frac{-n}{8 \log n})\exp\left[\frac{\Sigma_{i=1}^{n} P_i(\theta)(1 - P_i(\theta))}{64}\right.$$

$$\le 2\exp(\frac{-n}{16 \log n}) \ .$$

Combining (I.21) and (I.24) yields, recalling (I.14),

$$\text{(I.25)} \qquad \delta \equiv P[|Y - E(Y|\theta^{(n)})| \le s_n|\theta] \le 2\exp(\frac{-n}{16 \log n})$$

for θ satisfying (I.18). Recalling (I.17), it follows that

$$\delta \le \max\left(\exp(\frac{-x^2 n}{4}), 2\exp(\frac{-n}{16 \log n})\right)$$
$$\le 2\exp(\frac{-x^2 n}{4})$$

for θ satisfying $|\theta - \theta^{(n)}| \ge x'$ by the assumption that $x \le (\log n)^{-1/2}$

Hence, using (I.12),

$$P[|Y - \varphi^{(n)}| \le s_n, |\Theta - \theta^{(n)}| \ge x'] \le 2\exp(\frac{-x^2 n}{4}).$$

Thus, recalling (I.11) and (I.5), the desired result follows.

Let $k \equiv k_n$ denote a sequence of positive integers throughout the remainder of Section 4. Recall that

$$\text{(I.26)} \qquad Y_j^{(k)} \equiv \Sigma_{i=n+1}^{n+M} U_{ijk}/M$$

defines the assessment subtest score for examinee j of the partitioning interval $A_k^{(n)}$. It can easily be shown (and is intuitively clear) that

(I.27) $\{Y_j^{(k)}\}$ is identically distributed in j for fixed k and independent in j,k.

The proof, which follows from (2.3) and (2.4), is omitted. Let $Y_{j,P}$ be the partitioning subtest score of the jth examinee of the partitioning interval $A_k^{(n)}$. Recalling the notation of (2.1), let

(I.28) $$B_j \equiv B_{j,k}^{(n)} = [Y_{j,P} \in A_k^{(n)}], \alpha_i \equiv \alpha_{i,k}^{(n)} = P[U_{i+n,1k} = 1 | B_1]$$

and

(I.29) $$c \equiv c_n \equiv c_{n,k} = \sum_{i=1}^{M} \alpha_i(1 - \alpha_i)/M^2.$$

Denote the density of Θ_1 given B_1 by $f_n(\cdot) \equiv f_n^{(k)}(\cdot)$. Let $J \equiv J_k^{(n)}$ throughout.

In order to verify that the asymptotic distribution of $\hat{\sigma}_Y^2 - \hat{\sigma}_P^2 \equiv \hat{\sigma}_{Y,k}^2 - \hat{\sigma}_{P,k}^2$ has mean 0, it is necessary to consider the pre-asymptotic centering: .

<u>Proposition I.2</u>. Suppose (I.3) and that uniformly in $k(1 \le k \le K_n)$

(I.30) $$\frac{J(\log n)^4}{n^2} \to 0 \quad \text{as} \quad n \to \infty.$$

Then, letting $\sigma_Y^2 \equiv \sigma_{Y,n,k}^2 = V(Y_1^{(k)} | B_1)$, it follows uniformly in k that

(I.31) $$J^{1/2}(\sigma_Y^2 - c) \to 0 \quad \text{as } n \to \infty .$$

Suppose (I.3) and, letting $K \equiv K_n$, that uniformly in k,

(I.32) $$\frac{JK(\log n)^4}{n^2} \to 0 \quad \text{as } n \to \infty .$$

Then, uniformly in k ,

(I.33) $$(JK)^{1/2}(\sigma_Y^2 - c) \to 0 \quad \text{as } n \to \infty .$$

<u>Proof</u>. Only the proof of (I.33) will be given, since the proof of (I.31) is virtually identical. First, $\sigma_Y^2 = E(V(Y|\theta)) + V(E(Y|\theta))$. Thus ($\int$ denoting $\int_{-\infty}^{\infty}$)

$$\sigma_Y^2 = \sum_{i=n+1}^{n+M} \int P_i(\theta)(1 - P_i(\theta)) f^{(n)}(\theta) d\theta / M^2$$

$$+ \int \left(\frac{\sum_{i=n+1}^{n+M} P_i(\theta)}{M} - \frac{\sum_{i=n+1}^{n+M} \alpha_i}{M} \right)^2 f^{(n)}(\theta) d\theta$$

$$= \frac{\sum_{i=n+1}^{n+M} \alpha_i}{M^2} - \frac{\sum_{i=n+1}^{n+M} E(P_i^2(\Theta_1)|B_1)}{M^2} +$$

$$+ \int \left(\frac{\sum_{i=n+1}^{n+M} P_i(\theta)}{M} - \frac{\sum_{i=n+1}^{n+M} \alpha_i}{M} \right)^2 f^{(n)}(\theta)d\theta \quad .$$

Now,

$$\left(\frac{\sum_{i=n+1}^{n+M} P_i(\theta)}{M} - \frac{\sum_{i=n+1}^{n+M} \alpha_i}{M} \right)^2 = \left(\frac{\sum_{i=n+1}^{n+M} P_i(\theta)}{M} \right)^2$$

$$- \frac{2}{M^2} \sum_{i=n+1}^{n+M} P_i(\theta) \sum_{i=n+1}^{n+M} \alpha_i + \frac{\left(\sum_{i=n+1}^{n+M} \alpha_i \right)^2}{M^2}$$

$$= \frac{\sum_{i=n+1}^{n+M} P_i^2(\theta)}{M^2} + \frac{\sum_{n+1\le i\ne i'\le n+M} P_i(\theta)P_{i'}(\theta)}{M^2} - \frac{2}{M^2} \sum_{i=n+1}^{n+M} P_i(\theta) \sum_{i=n+1}^{n+M} \alpha_i$$

$$+ \frac{\sum_{i=n+1}^{n+M} \alpha_i^2}{M^2} + \frac{\sum_{n+1\le i\ne i'\le n+M} \alpha_i \alpha_{i'}}{M^2} \quad .$$

Thus,

$$\sigma_Y^2 - c = \sum_{n+1\le i\ne i'\le n+M} \frac{E[P_i(\Theta_1)P_{i'}(\Theta_1)|B_1]}{M^2}$$

$$- \frac{1}{M^2} \left(\sum_{i=n+1}^{n+M} \alpha_i \right)^2 + \frac{1}{M^2} \sum_{i=n+1}^{n+M} \sigma_i^2$$

$$= \frac{1}{M^2} \sum_{n+1\le i\ne i'\le n+M} (E[P_i(\Theta_1)P_{i'}(\Theta_1)|B_1] - \alpha_i\alpha_{i'}) \quad .$$

Thus, it suffices to show that, uniformly in $n + 1 \le i \ne i' \le n + M$,

$$(JK)^{1/2} \operatorname{cov}(P_i(\Theta_1), P_{i'}(\Theta_1)|B_1) \to 0 \quad .$$

Hence, it suffices to show

$$(JK)^{1/2} V(P_i(\Theta_1)|B_1) \to 0$$

as $n \to \infty$, uniformly in $n + 1 \le i \le n + M$.

Now, letting $x' = 2x/\varepsilon$

$$\text{(I.34)}\qquad V(P_i(\Theta_1)|B_1) = \int [P_i(\theta) - \int P_i(\theta')f^{(n)}(\theta')d\theta']^2 f^{(n)}(\theta)d\theta$$

$$= \int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} \left[P_i(\theta) - \int P_i(\theta')f^{(n)}(\theta')d\theta'\right]^2 f^{(n)}(\theta)d\theta$$

$$+ \int_{[\theta^{(n)}-x',\theta^{(n)}+x']^c} \left[P_i(\theta) - \int P_i(\theta')f^{(n)}(\theta')d\theta'\right]^2 f^{(n)}(\theta)d\theta$$

$$\equiv R + S \text{ , say .}$$

Now, by Proposition I.1, for any $s_n \le x \le (\log n)^{-1/2}$

$$\text{(I.35)}\qquad S \le \frac{C\exp(-nx^2/4)}{s_n(n^{1/2}s_n - n^{-1/2})} \equiv Q_n, \text{say} \quad .$$

Consider R. Then, using the mean value theorem, there exist θ_0 satisfying $|\theta_0 - \theta^{(n)}| \le x'$ such that

$$\text{(I.36)}\qquad |P_i(\theta) - \int P_i(\theta')f^{(n)}(\theta')d\theta'| =$$

$$|P_i(\theta) - \int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} P_i(\theta')f^{(n)}(\theta')d\theta' - \int_{[\theta^{(n)}-x',\,\theta^{(n)}+x']^c} P_i(\theta')f^{(n)}($$

$$\le \left| P_i(\theta) - \frac{\int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} P_i(\theta')f^{(n)}(\theta')d\theta'}{\int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} f^{(n)}(\theta')d\theta'} \int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} f^{(n)}(\theta')d\theta' \right| + Q_n$$

$$\le \left| P_i(\theta) - P_i(\theta_0)\int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} f^{(n)}(\theta')d\theta' \right| + Q_n \text{ ,}$$

$$= \left| P_i(\theta)\left[\int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} f^{(n)}(\theta')d\theta' + \int_{[\theta^{(n)}-x',\theta^{(n)}+x']^c} f^{(n)}(\theta')d\theta'\right] - P_i(\theta_0)\int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} f^{(n)}(\theta')d\theta' \right| + Q_n$$

$$\le |P_i(\theta) - P_i(\theta_0)|\int_{\theta^{(n)}-x'}^{\theta^{(n)}+x'} f^{(n)}(\theta)d\theta + 2\,Q_n \le Cx' + 2\,Q_n \quad ,$$

by the mean value theorem and (I.34). Thus, by (I.34), (I.35), and (I.36)

$$(I.37) \qquad (JK)^{1/2} V(P_i(\Theta_1)\,|\,B_1) \leq \left[(Cx' + Q_n)^2 + Q_n\right](JK)^{1/2} .$$

Thus, it suffices to obtain

$$(x')^2 (JK)^{1/2} \to 0 \quad , \quad Q_n (JK)^{1/2} \to 0 \quad .$$

Recall that $x' = 2x/\varepsilon$. Taking $x' = (\log n)/n^{1/2}$ works for example, establishing the result.

Next, the asymptotic distribution of $\hat{\sigma}_Y^2 - \hat{\sigma}_p^2$, suitably normalized, is obtained.

<u>Proposition I.3</u>. Suppose (I.3). Suppose for some $\varepsilon > 0$ and all $n + 1 \leq i \leq n + M$, $n \geq 1$ that (recalling that $\theta^{(n)} \equiv \theta_k^{(n)}$)

$$(I.38) \qquad |P_i(\theta^{(n)}) - 1/2| > \varepsilon .$$

Suppose (I.30). Let

$$(I.39) \qquad D \equiv D_k^{(n)} = (2\alpha_1 - 1,\ 2\alpha_2 - 1, \cdots, 2\alpha_M - 1, 1)$$

and, denoting $E(Y - E(Y))^4$ by $\mu_4(Y)$,

$$(I.40) \qquad \Sigma \equiv \Sigma_k^{(n)} = \begin{bmatrix} \alpha_1(1-\alpha_1) & 0 & \cdots & 0 & \alpha_1(1-\alpha)(1-2\alpha_1) \\ 0 & \alpha_2(1-\alpha_2) & \cdots & 0 & \alpha_2(1-\alpha_2)(1-2\alpha_2) \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \alpha_M(1-\alpha_M) & \alpha_M(1-\alpha_M)(1-2\alpha_M) \\ \alpha_1(1-\alpha_1)(1-2\alpha_1) & \alpha_2(1-\alpha_2)(1-2\alpha_2) & \cdots & \alpha_M(1-\alpha_M)(1-2\alpha_M) & M^4\left\{\mu_4(Y)-\sigma_Y^4\right\} \end{bmatrix}$$

Then, letting Z denote a $N(0,1)$ random variable,

$$(I.41) \quad M^2 \sqrt{J} \frac{\hat{\sigma}_Y^2 - \hat{\sigma}_P^2}{(D \Sigma D')^{1/2}} \xrightarrow{L} Z \text{ as } n \to \infty \text{, uniformly in } k .$$

Proof. Since writing "$\hat{\sigma}_Y^2 - \hat{\sigma}_P^2$" presupposes the assignment of examinees by the partitioning subtest, it is implicit in the statement of Proposition I.3 that the conclusion of (I.41) is conditional on $\bigcap_{j=1}^{J} B_j$.

Let, writing $Y_j \equiv Y_j^{(k)}$, $U_{ij} \equiv U_{n+i,jk}$, and $EY_j \equiv E[Y_j | B_j]$,

$$\underline{W}_j' = (U_{1j}, U_{2j}, \cdots, U_{Mj} \ , \ M^2(Y_j - EY_j)^2)$$

Note that

$$(I.42) \qquad E[\underline{W}'_1 | B_1] = (2\alpha_1 - 1, 2\alpha_2 - 1, \cdots, 2\alpha_M - 1 \ M^2 \sigma_Y^2) \ .$$

We next compute the asymptotic behavior of the covariance matrix $\Sigma_{\underline{W}}$ of $\underline{W}$ given B_1. Note that, by conditioning on Θ_1,

$$(I.43) \qquad \alpha_i = E(U_{i1} | B_1) = E[P_{n+i}(\Theta_1) | B_1]$$

and for $i \neq i'$

$$(I.44) \qquad E(U_{i1}, U_{i'1} | B_1) = E[P_{n+i}(\Theta_1) P_{n+i'}(\Theta_1) | B_1] \ .$$

Recalling the proof of Proposition I.2, for $i \neq i'$, it follows that

$$(I.45) \qquad J^{1/2} \operatorname{cov}(P_{n+i}(\Theta_1), P_{n+i'}(\Theta_1) | B_1) \to 0$$

and

$$J^{1/2} \ V(P_{n+i}(\Theta_1) | B_1) \to 0 \quad \text{as} \quad n \to \infty \ .$$

Thus, combining (I.43), (I.44), and (I.45), it follows that

$$(I.46) \qquad J^{1/2} \operatorname{cov}(U_{i1}, U_{i'1} | B_1) \to 0 \quad \text{as} \quad n \to \infty \ .$$

Now, by (4.7) and (I.2), for some $\varepsilon > 0$ and all i,k,n

(I.47) $$1 - \varepsilon \geq P_i(\theta^{(n)})(1 - P_i(\theta^{(n)})) \geq \varepsilon \ .$$

Now, for $x_n = (\log n)^{-1}$, using (4.4),

$$|\alpha_i - P_i(\theta^{(n)})| = \left|\int_{-\infty}^{\infty}(P_i(\theta) - P_i(\theta^{(n)}))f^{(n)}(\theta)d\theta\right|$$

$$\leq P[|\Theta_1 - \theta^{(n)}| > x_n|B_1] + Cx_n \ .$$

Hence, applying Proposition I.1, it follows that

(I.48) $$\alpha_i - P_i(\theta^{(n)}) \to 0 \quad \text{as} \quad n\to\infty$$

uniformly in i. Thus, combining (I.47) and (I.48), and recalling (I.29),

(I.49) $$\inf_{n\geq 1} c_n > 0 \ , \ 1 - \varepsilon \geq \alpha_i(1-\alpha_i) \geq \varepsilon \quad \text{for all } i \ .$$

Note that

(I.50) $$V(U_{i1}|B_1) = \alpha_i - \alpha_i^2 \ .$$

Further, using (I.46) and elementary calculation, denoting $a_n/b_n \to 1$ by $a_n \sim$

(I.51) $$\text{cov}(U_{i1}, M^2(Y_1-EY_1)^2|B_1) \sim E[(U_{i1}-\alpha_i)^3|B_1] = \alpha_i(1-\alpha_i)(1-2\alpha_i) \text{ as } n\to$$

Thus, recalling (I.40) and (I.46), (I.50), and (I.51), it has been establish that, termwise, as $n\to\infty$,

(I.52) $$\Sigma_{\underline{W}} \sim \Sigma \ .$$

Also, note by (I.38) and (I.48), it follows that, for some $\varepsilon > 0$

(I.53) $$|2\alpha_i - 1| > \varepsilon \ , \quad |\alpha_i(1-\alpha_i)(1-2\alpha_i)| > \varepsilon$$

uniformly in i.

Now, armed with the above computations, the proof proceeds essentially by a multivariate version of the " δ method" (see p. 122, Serfling (1980)). Let

(I.54) $$g(x_1,x_2,\cdots,x_{M+1}) = x_{M+1} - \Sigma_{i=1}^{M} x_i(1-x_i)$$

Thus

$$\frac{\partial g}{\partial x_i} = \begin{cases} 2x_i-1 & i \le M \\ 1 & i = M+1 \end{cases} .$$

Let, recalling (I.42),

$$\underline{X}_J = \Sigma_{j=1}^{J}\underline{W}_j/J \ , \quad \underline{\mu}_J = E[\underline{X}_J | \bigcap_{j=1}^{J} B_j] = E[\underline{W}_1 | B_1] \ .$$

The main thrust of the proof is the establishment of the asymptotic normality of $g(\underline{X}_J)$.

Let

$$h(\underline{x}) = \begin{cases} \dfrac{g(\underline{x}) - g(\underline{\mu}_J) - g(\underline{\mu}_J\ ;\ \underline{x} - \underline{\mu}_J)}{\|\underline{x} - \underline{\mu}_J\|} & \text{if } \ x \ne \underline{\mu}_J \\ 0 & \text{if } \ \underline{x} = \underline{\mu}_J \end{cases}$$

where

(I.55) $$g(\underline{\mu},\underline{t}) = \left.\frac{\partial g}{\partial x_1}\right|_{\underline{\mu}} t_1 + \cdots + \left.\frac{\partial g}{\partial x_{M+1}}\right|_{\underline{\mu}} t_{M+1} ,$$

the differential of g at $(\underline{\mu},\underline{t})$ and $\|\underline{x}\|^2 = \Sigma_1^{M+1} x_i^2$. Let

$$b_J = \frac{1}{J^{1/2}}(D\ \Sigma\ D')^{1/2} .$$

Now,

(I.56) $$\frac{g(\underline{X}_J) - g(\underline{\mu}_J)}{b_J} = h(\underline{X}_J)\frac{\|\underline{X}_J - \underline{\mu}_J\|}{b_J} + \frac{g(\underline{\mu}_J\ ;\ \underline{X}_J - \underline{\mu}_J)}{b_J} .$$

By (I.55), letting μ_{iJ} denote the ith component of $\underline{\mu}_J$,

$$\text{(I.57)} \qquad \frac{1}{b_J} g(\underline{\mu}_J;\underline{X}_J - \underline{\mu}_J) = \frac{1}{b_J} \sum_{i=1}^{M+1} (X_{iJ} - \mu_{iJ}) \left.\frac{\partial g}{\partial x_i}\right|_{\mu_i}$$

$$= \frac{1}{J^{1/2}(D\,\Sigma\,D')^{1/2}} \sum_{i=1}^{M} \left[\Sigma_{j=1}^{J} \left\{ (U_{ij} - \alpha_i)(2\alpha_i - 1) + M(Y_j - EY_j)^2 - \sigma_Y^2) \right\}\right.$$

$$= \frac{1}{J^{1/2}(D\,\Sigma\,D')^{1/2}} \sum_{j=1}^{J} \left[\Sigma_{i=1}^{M} (U_{ij} - \alpha_i)(2\alpha_i - 1) + M^2((Y_j - EY_j)^2 - \sigma_Y^2)\right.$$

Let

$$R_j \equiv R_j^{(k)} = \Sigma_{i=1}^{M} (U_{ij} - \alpha_i)(2\alpha_i - 1) + M^2((Y_j^{(k)} - EY_j^{(k)})^2 - \sigma_Y^2 \ ).$$

Now, using (I.52),

$$\text{(I.58)} \qquad V(R_j \mid \bigcap_{j'=1}^{J} B_{j'}) = V(R_j \mid B_j) \sim (D\,\Sigma\,D').$$

Moreover, by (2.3) and (2.4), it is easy to show that the $\{R_j\}$ are, for each n, i.i.d. By the Berry Esseen theorem, it suffices to show that

$$\inf_n \sigma_{R_1}^2 > 0,$$

which will follow from

$$\text{(I.59)} \qquad \inf_n (D\,\Sigma\,D')^{1/2} > 0$$

Since D is asymptotically bounded away from the $M+1$ dimensional vecto $\underline{0}$, it suffices to verify that asymptotically Σ is bounded away from the collection of $(M+1)\times(M+1)$ singular matrices. To this end, it is necessary to obtain an asymptotic expression for $\mu_4(Y) - \sigma_Y^4$. Computation sho that

$$\mu_4(Y) - \sigma_Y^4 \sim \Sigma_{i=1}^{M}(\alpha_i - 4\alpha_i^2 + 6\alpha_i^3 - 3\alpha_i^4)$$

$$+ 2 \sum_{1\le i\ne i'\le M} \alpha_i(1-\alpha_i)\alpha_{i'}(1-\alpha_{i'})$$

$$- \Sigma_{i=1}^{M} \alpha_i^2(1-\alpha_i)^2 \quad .$$

Then, although mildly tedious, using $|\alpha_i(1-\alpha_i)| > \varepsilon$ for all i, some $\varepsilon > 0$, it can be shown that Σ is indeed bounded away from the collection of singular matrices. Thus, noting (I.57) and applying the Lindeberg-Feller theorem, conditional on $\bigcap_{j=1}^{J} B_j$,

$$\text{(I.60)} \qquad \frac{1}{b_J}\, g(\underline{\mu}_J; X_J - \mu_J) \xrightarrow{L} Z ,$$

where Z is a $N(0,1)$ random variable. Now, since the elements of $\Sigma/J^{1/2}$ approach 0 as $n \to \infty$, it follows that, conditional on $\bigcap_{j=1}^{J} B_j$,

$$\underline{X}_J - \underline{\mu}_J \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty .$$

Hence, conditional on $\bigcap_{j=1}^{J} B_j$,

$$h(\underline{X}_J) - h(\underline{\mu}_J) \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty$$

by the continuity of h at $\underline{\mu}_J$. But, since $h(\underline{\mu}_J) = 0$, it follows that, conditional on $\bigcap_{j=1}^{J} B_j$,

$$h(\underline{X}_J) \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty .$$

Hence, referring to (I.56), to obtain asymptotic normality, it suffices to show that

$$P\left[\frac{\|X_J - \underline{\mu}_J\|}{b_J} > x \;\middle|\; \bigcap_{j=1}^{J} B_j\right] \to 0 \quad \text{as} \quad x \to \infty ,$$

uniformly in n. Hence, by Chebychev's inequality, it suffices that $D \Sigma D'$ is bounded below, which is known from (I.59). Thus, conditional on $\bigcap_{j=1}^{J} B_j$, letting Z denote a $N(0,1)$ random variable,

$$\text{(I.61)} \qquad \frac{g(\underline{X}_J) - g(\underline{\mu}_J)}{b_J} \xrightarrow{L} Z \quad \text{as} \quad n \to \infty \ .$$

Note that

$$\text{(I.62)} \qquad J^{1/2}(\hat{\sigma}_Y^2 - \sigma_Y^2) = J^{1/2}\left[\frac{\Sigma_{j=1}^{J}(Y_j - EY_j)^2}{J} - \sigma_Y^2\right] - J^{1/2}(\bar{Y} - EY_1)^2 \ .$$

Recall (I.49) and (I.31). Hence $\inf_n \sigma_Y^2 > 0$. Thus

$$\frac{n\ E(|Y_1 - EY_1|^3 | B_1)}{n^{3/2}\ \sigma_Y^2} \le \frac{C}{n^{1/2}} \ .$$

Hence by the Berry-Esseen theorem (Lemma I.1) it follows that, conditional on $\bigcap_{j=1}^{J} B_j$,

$$\text{(I.63)} \qquad J^{1/2}(\bar{Y} - EY_1) \xrightarrow{L} Z \ , \text{ a } N(0,1) \text{ random variable.}$$

By Chebychev's inequality, conditional on $\bigcap_{j=1}^{J} B_j$

$$\text{(I.64)} \qquad \bar{Y} - EY_1 \xrightarrow{P} 0 \ .$$

Thus, (I.63) and (I.64) together imply that, conditional on $\bigcap_{j=1}^{J} B_j$

$$J^{1/2}(\bar{Y} - EY_1)^2 \xrightarrow{P} 0 \ .$$

Thus, referring to (I.62) and (I.54), it follows that replacing $\Sigma_{j=1}^{J}(Y_j - EY_j)$ by σ_Y^2 in (I.61) leaves the conclusion of (I.61) unchanged. Applying (I.31) of Proposition I.2 to (I.61) and recalling (I.54) then yields the desired conclusion for each fixed $k = k_n$ sequence. Although not explicitly stated, all arguments in the above segment of the proof hold uniformly in k, thus establishing the desired conclusion.

Now the proof of Theorem 4.1 can be completed.

Proof of Theorem 4.1. Trivially follows from Proposition I.3 and the facts that conditional on $\bigcap_{j=1}^{J} B_j$,

(I.65) $$\hat{\alpha}_i - \alpha_i \xrightarrow{P} 0,\ \hat{\mu}_4(Y) - \mu_4(Y) \xrightarrow{P} 0,\ \hat{\sigma}_Y^2 - \sigma_Y^2 \xrightarrow{P} 0 \text{ as } n \to \infty$$

uniformly in i,k. Also note that conditioning on $\bigcap_{j=1}^{J} B_j$ is implicit in the statement of Theorem 4.1.

Remark. Note that, recalling (3.6) and (3.7)

(I.66) $$(\hat{D}\hat{\Sigma}\hat{D}') = M^4(\hat{\mu}_4(Y) - \hat{\sigma}^4_{Y,k}) - \hat{\delta}_{4,k}.$$

Now, the proof of Theorem 4.2 can be given.

Proof of Theorem 4.2. Note that condition on $B^{(n)} \equiv \bigcap_{k=1}^{K_k} \bigcap_{j=1}^{J_k^{(n)}} B_{j,k}^{(n)}$ is implicit in the statement of Theorem 4.2. The proof basically consists of a modification of the proof of Proposition I.3. The main steps are presented, with the details left to the interested reader: Refer to the proof of Proposition I.3. Let

$$\underline{W}_j^{(k)'} = (U_{n+1,jk}, U_{n+2,jk}, \ldots, U_{n+M,jk}, (Y_j^{(k)} - EY_j^{(k)})^2).$$

As in the derivation of (I.52), it follows that termwise, as $n \to \infty$,

(I.67) $$\Sigma_{\underline{W}^{(k)}} \sim \Sigma_k^{(n)}$$

uniformly on k. Define $g(\cdot)$ by (I.54). Let

$$\underline{X}_n = \sum_{k=1}^{K_n} \frac{\sum_{j=1}^{J_k^{(n)}} \underline{W}_j^{(k)}/J_k^{(n)}}{K_n},\quad \underline{\mu}_n = E[\underline{X}_n | B^{(n)}] = \sum_{k=1}^{K_n} E[\underline{W}_1^{(k)} | \bigcap_{k=1}^{K_n} B_1^{(k)}]/K_n.$$

Note that

$$g(\underline{X}_n) = \frac{1}{K_n} \sum_{k=1}^{K_n} \left[M^2 \frac{\sum_{j=1}^{J_k^{(n)}} (Y_j^{(k)} - EY_j^{(k)})^2}{J_k^{(n)}} - \sum_{k=1}^{M} \hat{p}_i^{(k)}(1 - \hat{p}_i^{(k)}) \right].$$

Define $g(\underline{\mu},\underline{t})$ and $h(\underline{x})$ as in the proof of Proposition I.3. Let

$$b_n^2 = \sum_{k=1}^{K_n} D_k^{(n)} \Sigma_k^{(n)} D_k^{(n)'} / J_k^{(n)}.$$

The rest of the proof proceeds analogously to that of Proposition I.3, noting that it is necessary to use (I.33) in place of (I.31). One concludes, conditional on $B^{(n)}$, that

$$\text{(I.68)} \qquad M^2 \frac{\sum_{k=1}^{K_n} \hat{\sigma}_{Y,k}^2 - \hat{\sigma}_{P,k}^2}{\left(\sum_{k=1}^{K_n} (D_k^{(n)} \Sigma_k^{(n)} D_k^{(n)'})/J_k^{(n)} \right)^{1/2}} \xrightarrow{L} Z.$$

Now, conditional on $B^{(n)}$,

$$\frac{\hat{D}_k^{(n)} \hat{\Sigma}_k^{(n)} \hat{D}_k^{(n)'}/J_k^{(n)}}{D_k^{(n)} \Sigma_k^{(n)} D_k^{(n)'}/J_k^{(n)}} \xrightarrow{P} 1 \quad \text{as} \quad n \to \infty$$

uniformly in k follows from the fact that (I.59) and (I.65) hold uniformly in k. Thus, noting again that (I.59) holds uniformly in k, conditional on $B^{(n)}$,

$$\text{(I.69)} \qquad \frac{\sum_{k=1}^{K_n} \hat{D}_k^{(n)} \hat{\Sigma}_k^{(n)} \hat{D}_k^{(n)'}/J_k^{(n)}}{\sum_{k=1}^{K_n} D_k^{(n)} \Sigma_k^{(n)} D_k^{(n)'}/J_k^{(n)}} \xrightarrow{P} 1 \quad \text{as} \quad n \to \infty.$$

Combining (I.68) and (I.69) yields the desired result.

Appendix II. Proof of Theorem 6.1 and Corollary 6.1.

Proof of Theorem 6.1. Note that conditioning on $B^{(n)}$ is implicit on the statement of (6.4). Let k denote the classical test dimensionality and R^k denote k dimensional Euclidean space. It follows trivially from Chebychev's inequality that the weak law of large numbers holds for the partitioning test score Y_1, uniformly in $\underline{\theta}$. That is,

$$\text{(II.1)} \qquad Y_1 - E[Y_1 | \underline{\Theta} = \underline{\theta}] \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty,$$

where the convergence in probability is uniform in $\underline{\theta} \in R^k$. It thus follows, denoting the distribution function of a random variable W by F_W and using (6.2), uniformly in $\underline{\theta}$ and y that

$$\text{(II.2)} \qquad F_{Y_1}(y \mid \underline{\Theta} = \underline{\theta}) - F_{E[Y_1 \mid \underline{\Theta}]}(y \mid \underline{\Theta} = \underline{\theta}) \to 0 \qquad \text{as} \qquad n \to \infty$$

for each $0 \leq y \leq 1$. Then, for fixed y, it is easy to show, using (6.2) and denoting the conditional density of $\underline{\Theta}$, given $Y = y$, by $f(\cdot \mid y)$ and the conditional density of $\underline{\Theta}$, given $E[Y \mid \underline{\Theta}] = y$ by $g(\cdot \mid y)$ that

$$\int_{R^k} \left[\frac{\sum_{i=1}^{M} P_i(\underline{\theta})}{M} - \int_{R^k} \frac{\sum_{i=1}^{M} P_i(\underline{\theta}')}{M} f(\underline{\theta}' \mid y) d\underline{\theta}' \right]^2 f(\underline{\theta} \mid y) d\underline{\theta}$$

$$- \int_{R^k} \left[\frac{\sum_{k=1}^{M} P_i(\underline{\theta})}{M} - \int_{R^k} \frac{\sum_{i=1}^{M} P_i(\underline{\theta}')}{M} g(\underline{\theta}' \mid y) d\theta' \right]^2 g(\theta \mid y) d\underline{\theta} \to 0$$

as $n \to \infty$, uniformly in y. Thus, integrating over y, it follows that

$$\text{(II.3)} \qquad \lim_{n \to \infty} \int_0^1 \int_{R^k} \left[\frac{\sum_{i=1}^{M} P_i(\underline{\theta})}{M} - \int_{R^k} \frac{\sum_{i=1}^{M} P_i(\underline{\theta}')}{M} f(\underline{\theta}' \mid y) d\underline{\theta}' \right]^2 f(\underline{\theta} \mid y) d\underline{\theta} \; dy$$

$$- \int_0^1 \int_{R^k} \left[\frac{\sum_{i=1}^{M} P_i(\underline{\theta})}{M} - \int_{R^k} \frac{\sum_{i=1}^{M} P_i(\underline{\theta}')}{M} g(\underline{\theta}' \mid y) d\underline{\theta} \right]^2 g(\underline{\theta} \mid y) d\underline{\theta} \; dy \to 0$$

as $n \to \infty$.

Similarly,

$$\text{(II.4)}\quad P_n - Q_n \equiv \frac{1}{K_n}\sum_{i=1}^{K_n}\int_{R^k}\left[\frac{\sum_{i=1}^{M} p_i(\underline{\theta})}{M} - \int_{R^k}\frac{\sum_{i=1}^{M} p_i(\underline{\theta}')}{M} f(\underline{\theta}' \mid Y_1 \in A_k^{(n)})d\underline{\theta}'\right]^2 f(\underline{\theta} \mid Y_1 \in A_k^{(n)})d\underline{\theta}$$

$$-\int_0^1\int_{R^k}\left[\frac{\sum_{k=1}^{M} p_i(\underline{\theta})}{M} - \int_{R^k}\frac{\sum_{i=1}^{M} p_i(\underline{\theta}')}{M} f(\underline{\theta}' \mid y)d\underline{\theta}'\right]^2 f(\underline{\theta} \mid y)d\underline{\theta}\, dy \to 0$$

as $n \to \infty$.

Combining (II.3), (II.4), and using (6.3), it thus follows that

$$\text{(II.5)}\quad \underline{\lim}_{n\to\infty} P_n > \frac{1+\varepsilon/2}{M}.$$

It is easy to see that, recalling (I.28) and (I.29),

$$\text{(II.6)}\quad E(\hat{\sigma}^2_{Y,k} \mid B^{(n)}) - E(\hat{\sigma}^2_{P,k} \mid B^{(n)}) = \frac{J_k^{(n)} - 1}{J_k^{(n)}}(\sigma^2_{Y,k} - c_{n,k}).$$

Arguing as in the proof of Proposition I.2,

$$\text{(II.7)}\quad \frac{1}{K_n}\sum_{k=1}^{K_n}(\sigma^2_{Y,k} - c_{n,k}) = \frac{-1}{K_n M^2}\sum_{k=1}^{K_n}\sum_{i=1}^{M} V[p_i(\underline{\Theta}) \mid Y_1 \in A_k^{(n)}] + P_n$$

$$\geq P_n - \frac{1}{M}.$$

Using (II.5), it thus follows that

$$\text{(II.8)}\quad \underline{\lim}_{n\to\infty}\frac{1}{K_n}\sum_{k=1}^{K_n}(\sigma^2_{Y,k} - c_{n,k}) \geq \frac{\varepsilon}{2M}.$$

Thus, using (II.6) and (II.7), for all n sufficiently large, conditional on $B^{(n)}$, it follows that

$$\text{(II.9)} \quad \frac{1}{K_n} \sum_{k=1}^{K_n} E(\hat{\sigma}^2_{Y,k} \mid B^{(n)}) - E(\hat{\sigma}^2_{P,k} \mid B^{(n)}) > \frac{\varepsilon}{2M} .$$

Now,

$$\text{(II.10)} \quad P[T_n > Z_\alpha \mid B^{(n)}] = P\Bigg[\frac{1}{K_n} \sum_{k=1}^{K_n} (\hat{\sigma}^2_{Y,k} - \hat{\sigma}^2_{p,k}) - \frac{1}{K_n} \{E(\hat{\sigma}^2_{Y,k} \mid B^{(n)}) - E(\hat{\sigma}^2_{P,k} \mid B^{(n)})\}$$

$$> M^2 \frac{Z_\alpha}{K_n} \left\{\sum_{k=1}^{K_n} (\hat{D}_k^{(n)} \hat{\Sigma}_k^{(n)} \hat{D}_k^{(n)'} / J_k^{(n)}\right\}^{1/2} - \frac{1}{K_n} \sum_{k=1}^{K_n} \Big\{E(\sigma^2_{Y,k} \mid B^{(n)}) - E(\hat{\sigma}^2_{P,k} \mid B^{(n)})\Big\} \mid B^{(n)}\Bigg]$$

But, trivially

$$\frac{\left\{\sum_{k=1}^{K_n} \hat{D}_k^{(n)} \hat{\Sigma}_k^{(n)} \hat{D}_k^{(n)'} / J_k^{(n)}\right\}^{1/2}}{K_n} \le C_n \to 0$$

as $n \to \infty$. Thus, using this and (II.9) and applying Chebychev's inequality, for large n

$$P[T_n > Z_\alpha \mid B^{(n)}] \ge 1 - \frac{C}{K_n} \to 1 \qquad \text{as} \qquad n \to \infty.$$

Thus, the result is proved.

Proof of Corollary 6.1. Note that conditioning on $Y_{1,P} \in A_k^{(n)}$ is implicit in the statement of Corollary 6.1. It is easy to show that

$$\sigma_{Y,k}^2 - c_{n,k} = 2\,\frac{\sum_{1\le i<i'\le M} \mathrm{cov}(U_i,U_{i'} \mid Y_{1,P} \in A_k^{(n)})}{M^2}$$

$$+\ \frac{\sum_{i=1}^{M} V(p_i(\underline{\Theta} \mid Y_{1,P} \in A_k^{(n)})}{M^2}.$$

Hence, (6.5) implies that

$$\varliminf_{n\to\infty} \frac{\sum_{k=1}^{K_n} (\sigma_{Y,k}^2 - c_{n,k})}{K_n} \ge \frac{\varepsilon}{2}\ ,$$

a result analogous to (II.8). The rest of the proof proceeds identically to that of Theorem 6.1 and is omitted.

References

Bartholomew, D.J., Factor analysis for categorical data, J. R. Stat. Soc., B., 293-321, (1980).

Chung, K.L., A Course in Probability Theory, Academic Press, New York, (1974).

Humphreys, L.C., General intelligence: an integration of factor, test, and simplex theory. To appear in a chapter of an upcoming memorial volumn in memory of David Wechsler, (1984).

Lamperti, J., Probability, W. A. Benjamin, New York, (1966).

Lawley, D.N. and Maxwell, A.E., Factor Analysis as a Statistical Method, 2nd. Ed., Elsevier, New York, (1971).

Lazersfeld, P.F. and Henry, N.W., Latent Structure Analysis, Wiley, New York, (1968).

Loève, M., Probability Theory I, 4th. Ed., Springer-Verlag, New York, (1977).

Lord, F.M. and Novick, M.R., Statistical Theories of Mental Test Scores, Addison-Wesley, Reading, Mass., (1968).

Lord, F.M., Applications of Item Response Theory to Practical Testing Problems, Lawrence Erlbaum, Hilsdale, N.J., (1980).

Sarrazin, G., The detection of item bias for different cultural groups using latent trait and chi-square methods. Preprint (1983).

Serfling, R.J., Approximation Theorems of Mathematical Statistics, Wiley, New York, (1980).

Solomon, H., A survey of mathematical models in factor analysis, In Mathematical Thinking in the Measurement of Behavior, Herbert Solomon, Editor, The Free Press, Glencoe, Illinois, 269-314, (1960).

Stout, W.F., Almost Sure Convergence, Academic Press, New York, (1974).

Navy

1 Dr. Nick Bond
Office of Naval Research
Liaison Office, Far East
APO San Francisco, CA 96503

1 Dr. Robert Breaux
NAVTRAEQUIPCEN
Code N-095R
Orlando, FL 32813

1 Dr. Robert Carroll
NAVOP 115
Washington , DC 20370

1 Dr. Stanley Collyer
Office of Naval Technology
800 N. Quincy Street
Arlington, VA 22217

1 CDR Mike Curran
Office of Naval Research
800 N. Quincy St.
Code 270
Arlington, VA 22217

1 Dr. John Ellis
Navy Personnel R&D Center
San Diego, CA 92252

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
Code P13
NPRDC
San Diego, CA 92152

1 Mr. Dick Hoshaw
NAVOP-135
Arlington Annex
Room 2834
Washington , DC 20350

1 Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

Navy

1 Dr William Montague
NPRDC Code 13
San Diego, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

1 Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

6 Personnel & Training Research Group
Code 442PT
Office of Naval Research
Arlington, VA 22217

1 Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

1 Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152

1 Mary Schratz
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Richard Snow
Liaison Scientist
Office of Naval Research
Branch Office, London
Box 39
FPO New York, NY 09510

1 Dr. Frederick Steinheiser
CNO - OP115
Navy Annex
Arlington, VA 20370

1 Mr. Brad Sympson
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Frank Vicino
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Edward Wegman
Office of Naval Research (Code 411S&P)
800 North Quincy Street
Arlington, VA 22217

Navy

1 Dr. Douglas Wetzel
Code 12,
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

1 Mr John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152

Marine Corps

1 Jerry Lehnus
CAT Project Office
HQ Marine Corps
Washington , DC 20380

1 Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

1 Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 Major Frank Yohannan, USMC
Headquarters, Marine Corps
(Code MPI-20)
Washington, DC 20380

Army

1 Dr. Kent Eaton
Army Research Institute
5001 Eisenhower Blvd.
Alexandria , VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Clessen Martin
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

1 Dr. William E. Nordbrock
FMC-ADCO Box 25
APO, NY 09710

1 Dr. Harold F. O'Neil, Jr.
Director, Training Research Lab
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Commander, U.S. Army Research Institute
for the Behavioral & Social Sciences
ATTN: PERI-BR (Dr. Judith Orasanu)
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Mr. Robert Ross
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joyce Shields
Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Hilda Wing
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Air Force

1 Dr. Alfred R. Fregly
AFOSR/NL
Bolling AFB, DC 20332

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Dr. Patrick Kyllonen
AFHRL/MOE

Brooks AFB, TX 78235

1 Lt. Col. Dave Payne
AFHRL/MOE
Brooks AFB, TX 78235

1 Dr. Roger Pennell
Air Force Human Resources Laboratory
Lowry AFB, CO 80230

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

1 Major John Welsh
AFHRL/MOAN
Brooks AFB , TX 78223

Department of Defense

12 Defense Technical Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defens
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 Dr. W. Steve Sellman
Office of the Assistant Secretary
of Defense (MRA & L)
2B269 The Pentagon
Washington, DC 20301

1 Dr. Robert A. Wisher
OUSDRE (ELS)
The Pentagon, Room 3D129
Washington, DC 20301

Civilian Agencies

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E Street NW
Washington, DC 20415

1 Mr. Thomas A. Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Private Sector

1 Dr. James Algina
University of Florida
Gainesville, FL 32605

1 Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Tel Aviv, Ramat Aviv 69978
Israel

1 Dr. R. Darrell Bock
Department of Education
University of Chicago
Chicago, IL 60637

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52243

1 Dr. Glenn Bryan
6208 Poe Road
Bethesda, MD 20817

1 Dr. John B. Carroll
409 Elliott Rd.
Chapel Hill, NC 27514

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. Hans Crombag
Education Research Center
University of Leyden
Boerhaavelaan 2
2334 EN Leyden
The NETHERLANDS

1 CTB/McGraw-Hill Library
2500 Garden Road
Monterey, CA 93940

Private Sector

1 Dr. Walter Cunningham
University of Miami
Department of Psychology
Gainesville, FL 32611

1 Dr. Dattpradad Divgi
Syracuse University
Department of Psychology
Syracuse, NE 33210

1 Dr. Emmanuel Donchin
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Dr. Hei-Ki Dong
Ball Foundation
Room 314, Building B
800 Roosevelt Road
Glen Ellyn, IL 60137

1 Dr. Fritz Drasgow
Department of Psychology
University of Illinois
603 E. Daniel St.
Champaign, IL 61820

1 Dr. Susan Embertson
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
Lawrence, KS 66045

1 ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Benjamin A. Fairbank, Jr.
McFann-Gray & Associates, Inc.
5825 Callaghan
Suite 225
San Antonio, TX 78228

1 Dr. Leonard Feldt
Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Univ. Prof. Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

Private Sector

1 Professor Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

1 Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01002

1 Dr. Robert Glaser
Learning Research & Development Center
University of Pittsburgh
3939 O'Hara Street
PITTSBURGH, PA 15260

1 Dr. Marvin D. Glock
217 Stone Hall
Cornell University
Ithaca, NY 14853

1 Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Ron Hambleton
School of Education
University of Massachusetts
Amherst, MA 01002

1 Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 90010

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
603 East Daniel Street
Champaign, IL 61820

1 Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

Private Sector

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Huynh Huynh
College of Education
University of South Carolina
Columbia, SC 29208

1 Dr. Douglas H. Jones
Advanced Statistical Technologies
Corporation
10 Trafalgar Court
Lawrenceville, NJ 08148

1 Dr. Marcel Just
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

1 Dr. Demetrios Karis
Department of Psychology
University of Illinois
603 E. Daniel Street
Champaign, IL 61820

1 Professor John A. Keats
Department of Psychology
The University of Newcastle
N.S.W. 2308
AUSTRALIA

1 Dr. William Koch
University of Texas-Austin
Measurement and Evaluation Center
Austin, TX 78703

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

1 Dr. Michael Levine
Department of Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Private Sector

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat 23
9712GC Groningen
Netherlands

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Mr. Phillip Livingston
Systems and Applied Sciences Corporatio
6811 Kenilworth Avenue
Riverdale, MD 20840

1 Dr. Robert Lockman
Center for Naval Analysis
200 North Beauregard St.
Alexandria, VA 22311

1 Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

1 Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

1 Dr. Don Lyon
P. O. Box 44
Higley , AZ 85236

1 Dr. Scott Maxwell
Department of Psychology
University of Notre Dame
Notre Dame, IN 46556

1 Dr. Samuel T. Mayo
Loyola University of Chicago
820 North Michigan Avenue
Chicago, IL 60611

1 Mr. Robert McKinley
American College Testing Programs
P.O. Box 168
Iowa City, IA 52243

Private Sector

1 Dr. Barbara Means
Human Resources Research Organization
300 North Washington
Alexandria, VA 22314

1 Dr. Robert Mislevy
711 Illinois Street
Geneva, IL 60134

1 Dr. W. Alan Nicewander
University of Oklahoma
Department of Psychology
Oklahoma City, OK 73069

1 Dr. Donald A Norman
Cognitive Science, C-015
Univ. of California, San Diego
La Jolla, CA 92093

1 Dr. James Olson
WICAT, Inc.
1875 South State Street
Orem, UT 84057

1 Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Cirle, NW
Washington, DC 20036

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 Dr. James W. Pellegrino
University of California,
Santa Barbara
Dept. of Psychology
Santa Barabara , CA 93106

1 Dr. Steven E. Poltrock
Bell Laboratories 2D-444
600 Mountain Ave.
Murray Hill, NJ 07974

1 Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Private Sector

1 Dr. Thomas Reynolds
University of Texas-Dallas
Marketing Department
P. O. Box 688
Richardson, TX 75080

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 Dr. Walter Schneider
Psychology Department
603 E. Daniel
Champaign, IL 61820

1 Lowell Schoer
Psychological & Quantitative
Foundations
College of Education
University of Iowa
Iowa City, IA 52242

1 Dr. Kazuo Shigemasu
7-9-24 Kugenuma-Kaigan
Fujusawa 251
JAPAN

1 Dr. William Sims
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

1 Dr. H. Wallace Sinaiko
Program Director
Manpower Research and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

1 Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Private Sector

1 Dr. Peter Stoloff
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Lab
252 Engineering Research Laboratory
Urbana, IL 61801

1 Dr. Maurice Tatsuoka
220 Education Bldg
1310 S. Sixth St.
Champaign, IL 61820

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Douglas Towne
Univ. of So. California
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

1 Dr. Robert Tsutakawa
Department of Statistics
University of Missouri
Columbia, MO 65201

1 Dr. Ledyard Tucker
University of Illinois
Department of Psychology
603 E. Daniel Street
Champaign, IL 61820

1 Dr. V. R. R. Uppuluri
Union Carbide Corporation
Nuclear Division
P. O. Box Y
Oak Ridge, TN 37830

Private Sector

1 Dr. David Vale
Assessment Systems Corporation
2233 University Avenue
Suite 310
St. Paul, MN 55114

1 Dr. Howard Wainer
Division of Psychological Studies
Educational Testing Service
Princeton, NJ 08540

1 Dr. Michael T. Waller
Department of Educational Psychology
University of Wisconsin--Milwaukee
Milwaukee, WI 53201

1 Dr. Brian Waters
HumRRO
300 North Washington
Alexandria, VA 22314

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 Dr. Rand R. Wilcox
University of Southern California
Department of Psychology
Los Angeles, CA 90007

1 Dr. Bruce Williams
Department of Educational Psychology
University of Illinois
Urbana, IL 61801

1 Ms. Marilyn Wingersky
Educational Testing Service
Princeton, NJ 08541

1 Dr. Wendy Yen
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940