SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): 1991-1	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: University of Illinois Department of Statistics; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Cognitive Science Program Office of Naval Research (Code 1142 CS)
6c. ADDRESS (City, State, and ZIP Code): 101 Illini Hall, 725 S. Wright Street, Champaign, IL 61820	7b. ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Arlington, VA 22217-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-87-K-0277
8c. ADDRESS (City, State, and ZIP Code): 101 Illini Hall, 725 S. Wright Street, Champaign, IL 61820	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61153N	RR04204	RR04204-01	4421-548

11. TITLE (Include Security Classification)
Essential Independence and Likelihood-based Ability Estimation for Polytomous Items

12. PERSONAL AUTHOR(S)
Brian W. Junker (Dept. of Statistics, Carnegie Mellon University, Pittsburgh PA 15213)

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
technical	FROM 1987 TO 1990	January 7, 1991	43

16. SUPPLEMENTARY NOTATION
To appear in Psychometrika, 1991

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	See reverse

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

See reverse

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL: Dr. Charles E. Davis	22b. TELEPHONE (Include Area Code): (703) 696-4046; 22c. OFFICE SYMBOL: ONR-1142-CS

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

Abstract

A definition of *essential independence* is proposed for sequences of polytomous items. For items satisfying the reasonable assumption that the expected amount of credit awarded increases with examinee ability, we develop a theory of *essential unidimensionality* which closely parallels that of Stout. Essentially unidimensional item sequences can be shown to have a unique (up to change-of-scale) dominant underlying trait, which can be consistently estimated by a monotone transformation of the sum of the item scores. In more general polytomous-response latent trait models (with or without ordered responses), an *M-estimator* based upon maximum likelihood may be shown to be consistent for θ under essentially unidimensional violations of local independence and a variety of monotonicity/identifiability conditions. A rigorous proof of this fact is given, and the standard error of the estimator is explored. These results suggest that ability estimation methods that rely on the summation form of the log-likelihood under local independence should generally be robust under essential independence, but standard errors may vary greatly from what is usually expected, depending on the degree of departure from local independence. An index of departure from local independence is also proposed.

KEY WORDS: item response theory (IRT), polytomous item responses, essential independence, unidimensionality, latent trait identifiability, likelihood-based trait estimation, asymptotic standard errors, structural robustness, local dependence.

1. Introduction

In the usual binary or dichotomous response formulation of item response theory (IRT), the correctness of the j^{th} item in a test or item sequence is indicated by a (random) response variable X_j taking on the value 1 for correct responses and the value 0 for incorrect responses. This codes the examinee's response with the score we wish to assign to that response. In considering polytomous data, it is convenient to treat the coding and scoring operations separately. For the j^{th} polytomous item we will code n possible response categories with the arbitrary labels $x_{j0}, x_{j1}, \ldots, x_{j(n-1)}$, and indicate the examinee's response with the (random) response variable

$$X_j \in \{x_{j0}, x_{j1}, \ldots, x_{j(n-1)}\}.$$

For convenience in scoring the item, it is also useful to have a set of *binary* response variables

$$Y_{jm} = \begin{cases} 1 & \text{if} \quad X_j = x_{jm}, \\ 0 & \text{else.} \end{cases}$$

Note that for each $j, Y_{j0} + Y_{j1} + \ldots + Y_{j(n-1)} = 1$, and that any item scoring method A_j that assigns the numerical score a_{jm} to the category x_{jm} may be expressed in terms of the $Y's$ as

$$A_j = \sum_{m=0}^{n-1} a_{jm} Y_{jm}.$$

Finally, let $\underset{\sim}{X}_J = (X_1, X_2, \ldots, X_J)$ be the vector of item responses on a test of length J given by a randomly-chosen examinee, and let $\underset{\sim}{x}_J = (x_1, x_2, \ldots x_J)$ denote any particular instance of $\underset{\sim}{X}_J$.

The general form of an IRT model for $\underset{\sim}{X}_J$ may then be expressed as

$$P[\, \underset{\sim}{X}_J = \underset{\sim}{x}_J \,] = \int P[\, \underset{\sim}{X}_J = \underset{\sim}{x}_J \mid \underset{\sim}{\Theta} = \underset{\sim}{\theta} \,] f(\underset{\sim}{\theta}) d\underset{\sim}{\theta}. \tag{1}$$

We follow Thissen and Steinberg (1986) in considering $\underset{\sim}{\Theta} = (\Theta_1, \ldots, \Theta_d)$, the latent trait or trait vector, to be a random variable (vector); thus, $f(\underset{\sim}{\theta})$ is this variable's probability density function for the population in question. The traditional IRT assumption of *local independence* reads, for polytomous item response models,

$$P[\, \underset{\sim}{X}_J = \underset{\sim}{x}_J \mid \underset{\sim}{\Theta} = \underset{\sim}{\theta}\,] = \prod_{j=1}^{J} \prod_{m=0}^{n-1} P_{jm}(\underset{\sim}{\theta})^{y_{jm}}, \tag{LI}$$

where the y_{jm} are observed values of Y_{jm} corresponding to each x_j, and $P_{jm}(\underset{\sim}{\theta}) \equiv P[X_j = x_{jm} | \underset{\sim}{\Theta} = \underset{\sim}{\theta}]$ are the response characteristic functions or, when $d = 1$, *response characteristic curves* (RCC's). There is no natural monotonicity assumption for general polytomous models, although for those cases in which the responses are ordered from least correct to most correct as m increases, it seems reasonable to require that

$$P^*_{jm}(\underset{\sim}{\theta}) \equiv \sum_{k=m}^{n-1} P_{jk}(\underset{\sim}{\theta}) \text{ is nondecreasing in } \underset{\sim}{\theta} \text{ for all } j, m, \tag{M}$$

that is, nondecreasing in each coordinate of $\underset{\sim}{\theta}$ with the other coordinates held fixed (these cumulative response functions are considered by, for example, Samejima, 1972). Note that $P^*_{jm}(\underset{\sim}{\theta}) = P[\text{response } m \text{ or greater } |\underset{\sim}{\theta}]$ is the binary item response function one would obtain by dichotomizing the item so that response m or greater is scored as 1 (correct) and any lower response is scored as 0 (incorrect). When LI and M both hold for a d-dimensional trait $\underset{\sim}{\Theta}$, we will write d_L for d. We will be concerned mostly with $d_L = 1$ models in what follows.

This paper has two aims. First, we wish to present and explore a definition of *essential independence* (EI) for polytomous item response sequences. EI, proposed for binary item sequences by Stout (1987; 1990), is a weakening of LI that is useful when—as seems often to be the case in real-life tests—there is a dominant underlying latent trait for the items but the presence of various minor traits prevents LI from holding exactly. For items satisfying a condition like M above, the theory of ***essential unidimensionality*** and estimation of the dominant unidimensional latent trait based on raw test score proceeds much as in Stout (1990). This is the subject of Sections 2 and 3.

Our second aim is to explore maximum likelihood estimates calculated under the assumption that LI holds when in fact only EI holds. Section 4 contains the basic result: the MLE calculated under LI remains consistent for θ under EI, subject only to regularity conditions and a natural identifiability condition. Thus, maximum likelihood estimation is robust against this realistic violation of local independence.

Monotone unidimensional local independence models will, and should, continue to be used as basic psychometric tools since they are attractive to the intuition and lead to explicit, analytically straightforward likelihoods. However, it is widely accepted that they oversimplify the latent structure of most tests in the real world. In some situations, the way the latent structure violates this simple model may be estimated and exploited, but in many situations it may be impossible or overly expensive to collect the data needed to ferret out a multidimensional latent structure. The discussion of this issue by Drasgow and Parsons (1983) is especially relevant here. Essential independence is a way of characterizing unidimensional stability without knowing the true likelihood function (latent structure). The importance of the robustness result of Section 4 is that it suggests that ability estimation methods based on the simple LI model continue to work in situations in which the latent factors causing strict LI to be violated are sufficiently minor that EI holds.

Despite this robustness in consistency, there is little robustness in variability. In

Section 5 we consider the standard error of the estimator of Section 4, showing that if the departure from local independence is great enough, the estimator can fail to have the usual standard error based on the information function, can fail to converge at the usual $J^{-1/2}$ rate, and can even fail to be asymptotically normally distributed. An index of the degree of departure from LI is proposed in Section 5 that can be used to calculate the new standard error. LI-based estimators like the MLE can be expected to be close to the examinee's θ under realistic conditions if the test is long, but conventional methods of assessing the standard errors of the estimates may be misleadingly optimistic in these same realistic settings.

Gibbons, Bock, and Hedeker (1989) have developed a method of factor analyzing dichotomous data with correlated specific factors that may be useful to obtain correct standard error estimates in at least some IRT settings. An indication of how their method might be used in the present context will be given in Section 5. Wainer and Wright (1980) have also reported some success using jackknife standard error estimates to account for extra variation in a $d_L = 1$ Rasch model due to guessing and "sleeping" behavior.

Also important in assessing the standard errors of ability estimators is the uncertainty involved in estimating RCC's. Tsutakawa and Soltys (1988) have incorporated RCC uncertainty into posterior mean estimator standard errors under LI in the dichotomous case. Adapting such methods to the EI setting will be of great importance in eventually understanding the true error structure of estimated IRT models, but that is beyond our present scope.

Although the results of this paper are stated and proved in the polytomous case, it is expected that they will find greatest application in the dichotomous setting, where IRT techniques have been most fully developed. For the reader's convenience, the main points of Sections 4 and 5 are restated for dichotomous responses in Sections 6—these results are also new in the dichotomous case. Finally, Section 7 summarizes the conclusions of our work, and indicates extensions to other popular LI-based trait estimators, such as the posterior mode and posterior mean.

2. Essential Independence and Item Sequences

The notions of *essential independence* and *essential unidimensionality* were introduced in Stout (1987) and explored in the dichotomous case by Stout (1990) and Junker (1988). In the factor analytic tradition, but with a decidedly non-factor-analytic perspective, Stout seeks a criterion by which only *dominant* dimensions can be counted. When only one dominant dimension is counted, the test is said to be *essentially unidimensional.*

The fundamental idea behind essential independence is that a trait vector $\underset{\sim}{\Theta}$ is dominant if, after conditioning on $\underset{\sim}{\Theta}$, the residual covariances among the items are small on average. This parallels the idea, in traditional IRT, that if the latent space is "complete", then the residual covariances are all zero. A partial answer to the question of how small the residual covariances must be for $\underset{\sim}{\Theta}$ to dominate has been provided by Stout's (1987) statistical procedure for assessing essential unidimensionality in a fixed, finite set of dichotomous items. If the residual covariances are small but not zero, $\underset{\sim}{\Theta}$ continues to have many properties of LI latent trait vectors: it is strongly related to the total test score, it is better and better identified as the test length grows, etc.

To examine properties of $\underset{\sim}{\Theta}$ and of $\underset{\sim}{\theta}$ estimators as test length grows, it is necessary to embed the finite test $X_1, \ldots, X_J$ in an infinite collection of items $\underset{\sim}{X}$. For example, results of Levine (1989) make it clear that not even the distribution of $\underset{\sim}{\Theta}$ is completely identifiable from a finite-length test, let alone particular examinees' $\underset{\sim}{\theta}$ vectors. Such an embedding is implicit even in traditional discussions of IRT trait estimation (e.g., Birnbaum, 1968, pp. 455–457; Lord, 1980, p. 59).

The substantive interpretation of this embedding varies from application to application. In some settings it may be reasonable to imagine that the process used to generate the test $X_1, \ldots, X_J$—which may, for example, involve many item writers and reviewers generating items of the same character and in the same way—is simply continued to produce more and more items. Or, it may be reasonable to think of $X_1, \ldots, X_J$ as forming a (stratified) sample from a large item pool, as when test forms are constructed by hand according to a test specification matrix, or constructed "on the fly" in computerized adaptive testing (CAT). Other interpretations may also be appropriate.

All such interpretations may be encompassed in the following framework. In practice, a test form of length $J + 1$ is seldom obtained by simply finding a form of length J and tacking one more item onto the end of it. Instead, forms of differing lengths—intended to measure the same construct—will be constructed at different times according to slightly different design specifications. Thus, in attempting to understand what is meant by letting the test length J grow, we may consider a sequence of *tests*

$$\begin{aligned}
\underset{\sim}{X}_1 &= (X_{11}), \\
\underset{\sim}{X}_2 &= (X_{21}, X_{22}), \\
\underset{\sim}{X}_3 &= (X_{31}, X_{32}, X_{33}), \\
&\vdots \\
\underset{\sim}{X}_J &= (X_{J1}, X_{J2}, X_{J3}, \cdots, X_{JJ}), \\
&\vdots
\end{aligned}$$

in which the test of length J need *not* be a subtest of the test of length $J+1$, for any J. The only requirement here is that each test be designed to measure the same construct. LI and other properties of the traditional IRT model extend in a natural way to such a sequence of tests by requiring that they hold in every test $\underset{\sim}{X}_J$ in the sequence. We will abstract the idea that the tests "measure the same construct" by assuming that $\underset{\sim}{\Theta}$ is the same from test to test, and that when an item appears in more than one $\underset{\sim}{X}_J$, it has the same response curves each time it appears.

This framework allows us to make mathematically rigorous statements about the identifiability, uniqueness, and estimation of dominant latent traits as test length grows. It is justifiable insofar as it helps crystalize ideas about finite-length tests with both dominant and minor dimensions, or it suggests ways to improve the analysis of real tests. The sense in which $\underset{\sim}{\Theta}$ is the dominant influence, essential independence, will be carefully defined in the next section. For now we remark that it is not necessary to arrange the items within $\underset{\sim}{X}_J$ in any particular order to acheive this. Rather, essential independence requires that the relative influence of minor factors not included in $\underset{\sim}{\Theta}$ be weaker—through cancellation between items, moderation within items, etc.—in longer tests than in shorter ones.

Formally, this framework leads to a rather messy notation, since it adds a "test index" J to all quantities under discussion: a_{jm} becomes a_{Jjm}, A_j becomes A_{Jj}, etc. For simplicity's sake, we will retain the notation of Section 1 in what follows, and speak informally of embedding the fixed test $\underset{\sim}{X}_J$ as the first J items in a single infinite item sequence $\underset{\sim}{X} = (X_1, X_2, X_3, \ldots)$. The reader should bear in mind that the results below also apply to the more general framework described above.

3. Essential Independence for Polytomous Items

The traditional approach in IRT is to say that a latent trait (vector) $\underset{\sim}{\Theta}$ completely controls the interesting variation in the item responses if LI and M hold. In contrast, we would like to be able to determine whether the latent vector $\underset{\sim}{\Theta}$ is the dominant influence underlying the item responses. Moreover, $\underset{\sim}{\Theta}$ should dominate regardless of how the responses are scored. Thus, it is appropriate to consider an arbitrary scoring scheme $\{a_{jm}\}$ and corresponding item scores A_j subject only to the constraint that there is some $M < \infty$ such that $|a_{jm}| \leq M$ for all j, m. All of the scoring schemes considered below will be *bounded* in this manner. If $\underset{\sim}{\Theta}$ is to be the dominant latent trait vector, we should at least require that the variation of the raw score, $\bar{A}_J = \frac{1}{J}\sum_{j=1}^{J} A_j$, be small when we condition on $\underset{\sim}{\Theta}$, as $J \to \infty$.

Definition 3.1. The sequence of polytomous items $\underset{\sim}{X}$ is ***essentially independent*** (EI) with respect to the latent trait(s) $\underset{\sim}{\Theta}$ if and only if, for every bounded scoring scheme $\{a_{jm}\}$ and every $\underset{\sim}{\theta}$,

$$\lim_{J\to\infty} \binom{J}{2}^{-1} \sum_{i=1}^{J}\sum_{j=1}^{i-1} \mathrm{Cov}(A_i, A_j \mid \underset{\sim}{\Theta} = \underset{\sim}{\theta}) = 0. \tag{EI}$$

This definition of EI for polytomous items, which is equivalent to requiring that $\lim_{J\to\infty} \mathrm{Var}(\bar{A}_J \mid \underset{\sim}{\Theta} = \underset{\sim}{\theta}) = 0$ for every bounded scoring scheme, directly generalizes Stout's definition of *strong EI* for binary items (Definition 3.5, Stout, 1990). Stout's various definitions of essential independence are likely not equivalent in general, but they are equivalent when the residual covariances are nonnegative (as seems plausible in many educational testing contexts; see the discussion following Theorem 5.1 below). Only the strong EI definition generalizes naturally to the polytomous case, and for this reason it is preferred in this paper.

Clearly, every LI item sequence is EI. Since the covariances above are unaffected by shifting the coefficients from a_{jm} to $a'_{jm} = a_{jm} + c_j$, for any constants c_j, we see that Definition 3.1 is equivalent to ones in which only positive, bounded a_{jm} are allowed; or only bounded a_{jm} for which at least one response from each item has $a_{jm} = 0$ are allowed; etc. Now, consider the expected item scores,

$$A_j(\underset{\sim}{\theta}) = E[A_j \mid \underset{\sim}{\Theta} = \underset{\sim}{\theta}] = \sum_{m=0}^{n-1} a_{jm} P_{jm}(\underset{\sim}{\theta}),$$

and the expected raw test score, or test characteristic function,

$$\bar{A}_J(\underset{\sim}{\theta}) = \frac{1}{J} \sum_{j=1}^{J} A_j(\underset{\sim}{\theta}).$$

Theorem 3.1. The following are equivalent, for a sequence of polytomous items $\underset{\sim}{X}$:

(a) $\underset{\sim}{X}$ is EI with respect to $\underset{\sim}{\Theta}$;

(b) For each bounded scoring scheme $\{a_{jm}\}$ and each $\underset{\sim}{\theta}$,

$$\lim_{J \to \infty} E\left[\, (\bar{A}_J - \bar{A}_J(\underset{\sim}{\theta}))^2 \mid \underset{\sim}{\Theta} = \underset{\sim}{\theta} \,\right] = 0;$$

(c) For each bounded scoring scheme $\{a_{jm}\}$ and each $\underset{\sim}{\theta}$,

$$\frac{1}{J} \sum_{j=1}^{J} \sum_{m=0}^{n-1} a_{jm} \left[\, Y_{jm} - P_{jm}(\underset{\sim}{\theta}) \,\right] \to 0$$

in probability, given $\underset{\sim}{\Theta} = \underset{\sim}{\theta}$, as $J \to \infty$.

Proof: The proof is an easy extension of the proof of Theorem 3.2 of Stout (1990). □

Estimating $\bar{A}_J(\underset{\sim}{\theta})$ is not necessarily useful unless $\underset{\sim}{\Theta}$ is unidimensional. Just as with binary items, a particular value $\bar{A}_J(\underset{\sim}{\theta})$ may be possible for examinees with radically different $\underset{\sim}{\theta}$'s due to compensation among the components of $\underset{\sim}{\theta}$. Hereafter, we will restrict ourselves to unidimensional traits Θ and consider estimating each examinee's θ.

When Θ is unidimensional, some sort of monotonicity condition becomes useful, so that we can estimate θ with $\tilde{\theta}_J = \bar{A}_J^{-1}(\bar{A}_J)$, where $\bar{A}_J^{-1}(\cdot)$ is the inverse of $\bar{A}_J(\theta)$. (In the usual binary setting $\bar{A}_J^{-1}(\bar{A}_J) = \bar{P}_J^{-1}(\bar{X}_J)$, for example.) In models that award partial credit for partially-correct answers, it seems natural to require that the expected amount of partial credit awarded on each item increases with the level of the latent trait:

$$A_j(\theta) \text{ is nondecreasing in } \theta \text{ for each } j. \qquad (\mathrm{M}')$$

What is the relationship between condition M in Section 1 and M′ above? We will call a sequence of items $\underset{\sim}{X}$ for which the item response categories $\{x_{jm}\}$ are indexed so that M holds an *ordered-response item sequence*. On the other hand, if a scoring scheme $\{a_{jm}\}$ satisfies, for each $j, 0 \leq a_{j0} \leq a_{j1} \leq \cdots \leq a_{j(n-1)}$, we will call it a *ordered scoring scheme*. Then, with the convention that $a_{j(-1)} \equiv 0$,

$$A_j(\theta) = \sum_{k=0}^{n-1} a_{jk} P_{jk}(\theta) = \sum_{m=0}^{n-1} (a_{jm} - a_{j(m-1)}) P^*_{jm}(\theta).$$

It follows that *condition* M *is equivalent to* M′ *holding for every ordered-response scoring scheme.* M is a condition that has been considered for many parametric ordered-response models. For example, Samejima (1972) has shown that M does hold for her graded-response model, as well as for Bock's (1972) nominal model constrained to apply to ordered-response items (also, see Thissen and Steinberg, 1986). A somewhat milder form of monotonicity called LAD is sufficient to build the estimator $\bar{\theta}_J$.

Definition 3.2. The ordered scoring scheme $\{a_{jm}\}$ is *asymptotically discriminating* (AD) if and only if there exists an $\epsilon > 0$ such that

$$\frac{1}{J}\sum_{j=1}^{J}(a_{j(n-1)} - a_{j0}) \geq \epsilon, \ \forall \ J. \tag{AD}$$

The item sequence $\underset{\sim}{X}$ is *locally asymptotically discriminating* (LAD) if and only if, for each AD ordered scoring scheme $\{a_{jm}\}$, to every θ there corresponds an interval N_θ containing θ and an $\epsilon_\theta > 0$ such that

$$\frac{\bar{A}_J(t) - \bar{A}_J(\theta)}{t - \theta} \geq \epsilon_\theta, \ \forall \ t \in N_\theta, \ t \neq \theta, \ \forall \ J. \tag{LAD}$$

This generalizes LAD for binary item sequences as presented in Definition 3.8 of Stout (1990). Note that LAD imposes a minimum discrimination condition on the test characteristic curves at each θ, as $J \to \infty$. Also, the items themselves need not have ordered responses; only the scoring schemes $\{a_{jm}\}$ need be ordered. LAD may be viewed as naturally extending the interpretation of M—that the expected amount of credit awarded increases with the examinee's ability—from a fixed-length test to an item sequence, without strictly requiring M to hold for every item in the sequence.

Theorem 3.2. If the polytomous item sequence $\underset{\sim}{X}$ satisfies EI and LAD with respect to the unidimensional trait Θ, then for each θ and each $\epsilon > 0$, if $\{a_{jm}\}$ is a bounded AD ordered scoring scheme, then

$$\lim_{J\to\infty} P\,[\,|\,\bar{A}_J^{-1}(\bar{A}_J) - \theta\,| > \epsilon \,|\, \Theta = \theta\,] = 0.$$

Proof: Virtually the same as the proof of Theorem 3.6 of Stout (1990). □

Theorem 3.3. If the polytomous item sequence $\underset{\sim}{X}$ satisfies EI and LAD with respect to the unidimensional trait Θ, and satisfies EI with respect to another latent trait τ, then there exists a nondecreasing function $g(t)$ such that

$$P[\,\Theta = g(\tau)\,] = 1.$$

Proof: Follows Theorem 3.3 of Stout (1990) or Theorem 2.4 of Junker (1988). □

Theorems 3.1 through 3.3 show that if EI and LAD hold, we can estimate a unique dominant latent trait with any reasonable $\bar{A}_J^{-1}(\bar{A}_J)$: any other dominant trait we might find will be change-of-scale of the trait we have estimated with $\bar{A}_J^{-1}(\bar{A}_J)$. (This is the same level of trait uniqueness as exists under the general $d_L = 1$ model, although particular parametic models—for example, the Rasch model—may possess additional scale properties.) Since under EI and LAD we can identify and estimate a unique unidimensional dominant trait in the item response data, we will call this situation *essentially unidimensional*, $d_E = 1$. When no single dominant trait exists in this sense, we will write $d_E > 1$.

4. Maximum Likelihood Ability Estimation

Often it is desired to estimate individuals' θ values, treated as parameters in the conditional model.

$$P[\underset{\sim}{X}_J = \underset{\sim}{x}_J \mid \Theta = \theta] = \prod_{j=1}^{J} \prod_{m=0}^{n-1} P_{jm}(\theta)^{y_{jm}},$$

where $y_{jm} = 1$ when $x_j = x_{jm}$, and 0 otherwise (i.e., y_{jm} are the observed values of Y_{jm}). If the polytomous item sequence $\underset{\sim}{X}$ does not satisfy LAD, the estimators described in Section 3 may not exist, let alone be consistent for θ. Even when LAD holds, it may be desirable to have a more-efficient estimator than $\bar{A}_J^{-1}(\bar{A}_J)$.

One common method of estimating individual examinees' abilities is via *maximum likelihood*, treating each examinee's θ as an unknown parameter to be estimated and the RCC's as known. When LI holds, the maximum likelihood estimator (MLE) $\hat{\theta}_J$ is known to be a consistent estimator for θ as $J \to \infty$, and has good asymptotic distribution properties (asymptotic normality, efficiency, etc.), assuming that the RCC's are known (e.g., Lehmann, 1983). We wish to investigate the behavior of $\hat{\theta}_J$ computed under the (false) assumption of LI with respect to Θ when the item sequence satisfies EI. Technically, $\hat{\theta}_J$ is called an *M-estimator* since it is no longer based on the true likelihood, which is unknown under EI (e.g., Serfling, 1980, pp. 243 ff.). However, for convenience we will continue to call $\hat{\theta}_J$ the MLE, since it is based on maximizing a (wrong) likelihood.

There are two reasons for working with the MLE. First, it is commonly used for examinee scoring in applied IRT work, so we are compelled to know its behavior under realistic violations of LI. Second, the behavior of the MLE may be taken to be representative of the behavior of other likelihood-based methods. Our work with the MLE is intended to suggest that similar robustness to departures from LI within an EI framework could be expected of other popular estimators and predictors, such as estimators of the posterior mode and posterior mean (e.g., Samejima, 1969; Bock & Mislevy, 1982; Lord, 1986). This point will be taken up again in the discussion in Section 7.

Let us now turn to the requirements for consistency of $\hat{\theta}_J$, the convergence of $\hat{\theta}_J$ to θ

as J grows. Assuming (incorrectly) that LI holds with respect to Θ, the log-likelihood for estimating one examinee's θ based on his or her item responses $\underset{\sim}{X}_J$—or equivalently, the response-category indicators Y_{jm}—is

$$\ell_J(\theta) = \log \prod_{j=1}^{J} \prod_{m=0}^{n-1} P_{jm}(\theta)^{Y_{jm}} = \sum_{j=1}^{J} \sum_{m=0}^{n-1} \lambda_{jm}(\theta) Y_{jm},$$

where $\lambda_{jm}(\theta) = \log P_{jm}(\theta)$. Thus, $\hat{\theta}_J$ must satisfy the likelihood equation

$$\frac{1}{J}\ell'_J(\hat{\theta}_J) = \frac{1}{J}\sum_{j=1}^{J} \sum_{m=0}^{n-1} \lambda'_{jm}(\hat{\theta}_J) Y_{jm} = 0. \tag{2}$$

Under LI, the fact that $\frac{1}{J}\ell'_J(\theta) \approx 0$ as $J \to \infty$ allows us to locate a root $\hat{\theta}_J$ of (2) near the examinee's θ. Under EI, Theorem 3.1(c) ensures that

$$\frac{1}{J}\ell'_J(\theta) = \frac{1}{J}\sum_{j=1}^{J} \sum_{m=0}^{n-1} \lambda'_{jm}(\theta)[Y_{jm} - P_{jm}(\theta)] \to 0, \tag{3}$$

in probability, given $\Theta = \theta$, as long as the scoring scheme $a_{jm} = \lambda'_{jm}(\theta)$ is bounded uniformly in j and m for each θ. Hence, we can expect to find a root $\hat{\theta}_J$ of (2) near θ under EI as well. The dependence of a_{jm} on θ here is irrelevant, since we are conditioning on $\Theta = \theta$ fixed.

To obtain the limit (3) and similar limits needed for consistency of $\hat{\theta}_J$, we assume that for all θ, there exists an interval B_θ containing θ and a constant $M_\theta < \infty$, such that

$$|\lambda'''_{jm}(t)| \leq M_\theta \ \forall\, t \in B_\theta,\ \forall\, j,\ m. \tag{4}$$

Condition (4) is really a fairly mild modeling assumption. For example, in the binary three parameter logistic model it would be satisfied if all the difficulty and discrimination parameters were bounded in absolute value.

A second important consideration in likelihood-based (or indeed any) estimation is identifiability of the parameter. The criterion used for identifiability in Section 3, LAD, is not necessarily appropriate when the response categories are unordered. Instead, it is typical and reasonable to require that for each θ, there exists an $\epsilon_\theta > 0$ such that

$$\bar{I}_J(\theta) = \frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1} \lambda'_{jm}(\theta)P'_{jm}(\theta) \geq \epsilon_\theta, \ \forall\ J. \tag{5}$$

$\bar{I}_J(\theta)$ is, of course, the usual *test information function*. If (5) holds, there is enough identifiability for the MLE to work. The following proposition gives several sufficient criteria for identifiability in this sense.

Proposition 4.1. If any of the following conditions hold, then (5) holds.

(a) $\forall\ \theta,\ \exists\ \epsilon_\theta > 0 : \frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1}\frac{[P'_{jm}(\theta)]^2}{P_{jm}(\theta)} \geq \epsilon_\theta,\ \forall\ J$;

(b) $\forall\ \theta,\ \exists\ \epsilon_\theta > 0 : \frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1}[P'_{jm}(\theta)]^2 \geq \epsilon_\theta,\ \forall\ J$;

(c) $\forall\ \theta,\ \exists\ \epsilon_\theta > 0 : \frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1}\mid P'_{jm}(\theta)\mid \geq \epsilon_\theta,\ \forall\ J$;

(d) $\forall\ \theta,\ \exists\ \epsilon_\theta > 0 : \frac{1}{J}\sum_{j=1}^{J}\sum_{m=1}^{n-1} P'_{jm}(\theta) \geq \epsilon_\theta,\ \forall\ J$.

Proof: Condition (a) is exactly (5); condition (b) suffices by (a) and the fact that $1/P_{jm}(\theta) > 1$ always holds; condition (c) suffices by (b) and the fact that

$$\frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1}[P'_{jm}(\theta)]^2 \geq \frac{1}{n}\left\{\frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1}\mid P'_{jm}(\theta)\mid\right\}^2,$$

by Jensen's inequality (Ash, 1972, p. 287). Finally, condition (d) suffices by noting that

$$\bar{I}_J(\theta) = \frac{1}{J}\sum_{j=1}^{J}\left\{\sum_{m=1}^{n-1}\frac{[P'_{jm}(\theta)]^2}{P_{jm}(\theta)} + \frac{1}{P_{j0}(\theta)}\left\{\sum_{m=1}^{n-1}P'_{jm}(\theta)\right\}^2\right\},$$

and, using Jensen's inequality again,

$$\frac{1}{J}\sum_{j=1}^{J}\frac{1}{P_{j0}(\theta)}\left\{\sum_{m=1}^{n-1}P'_{jm}(\theta)\right\}^2 \geq \left\{\frac{1}{J}\sum_{j=1}^{J}\frac{1}{P_{j0}(\theta)}\sum_{m=1}^{n-1}P'_{jm}(\theta)\right\}^2. \ \square$$

The most interesting of the criteria in Proposition 4.1 is (d). Note that by taking $a_{j0} = 0$ and $a_{jm} \equiv 1$ for $m > 0$ in the definition of LAD, we see that if LAD holds and the RCC's are differentiable, then (5) holds also.

Each of the conditions M, LAD, and (5) represent identifiability or detection conditions for the sequence $\underset{\sim}{X}$ and latent trait Θ, and they fit into a rather neat hierarchy for essentially independent smooth IRT models. M is the most restrictive identification condition; it imposes a highly interpretable condition on each item in the test which virtually guarantees LAD. LAD is less restrictive, in that it imposes the interpretation of M at the test characteristic curve level, not the level of individual items. Moreover, LAD implies (5). The minimum information condition (5) is least interpretable, but has the advantage of widest applicability. Moreover, as Theorem 4.1 below shows, if (5) holds, then $\hat{\theta}_J$ converges to θ, given $\Theta = \theta$. This hierarchy is not new or deep mathematically, but serves to illustrate the transition from intuitively appealing psychological models to adequate but less pleasing statistical ones.

Theorem 4.1. Let $\underset{\sim}{X}$ be a polytomous item sequence satisfying EI, (4) and (5). Then there exists a sequence $\{\hat{\theta}_J : J \geq J_\theta\}$ of roots of (2) such that

$$\lim_{J \to \infty} P[\mid \hat{\theta}_J - \theta \mid < \epsilon \big| \Theta = \theta] = 1,$$

for every $\epsilon > 0$.

Note that the sequence $\hat{\theta}_J$ may not start at $J = 1$, and for small J, there may be no solutions to (2). This is not a serious limitation; see Theorem 4.2 below. Also, when LAD holds, the trait being estimated is the same dominant trait whose estimation was treated in Section 3; this follows from Theorem 3.3. The novelty of the following Cramér-style proof is that (local) independence is not assumed.

Proof: Let $\epsilon_0 > 0$ be arbitrary and fixed in advance. Without loss of generality, we assume that $(\theta - \epsilon_0, \theta + \epsilon_0) \subset B_\theta$, where B_θ is the interval given in (4). Our goal is to obtain roots of (2) in the interval $(\theta - \epsilon_0, \theta + \epsilon_0)$. The second-order Taylor polynomial for $\frac{1}{J}\ell'_J(t)$ in $(\theta - \epsilon_\theta, \theta + \epsilon_\theta)$ is

$$\frac{1}{J}\ell'_J(t) = \frac{1}{J}\ell'_J(\theta) + \frac{1}{J}(t-\theta)\ell''_J(\theta) + \frac{1}{2J}(t-\theta)^2\ell'''_J(\xi) \tag{6}$$

$$= \frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1} \lambda'_{jm}(\theta)Y_{jm}$$

$$+(t-\theta)\frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1} \lambda''_{jm}(\theta)Y_{jm} + (t-\theta)^2\frac{1}{2J}\sum_{j=1}^{J}\sum_{m=0}^{n-1} \lambda'''_{jm}(\xi)Y_{jm},$$

where $\xi = \theta + r(t-\theta)$, for some $0 \le r \le 1$. We have already shown in (3) that $\frac{1}{J}\ell'_J(\theta) \to 0$ in probability, given $\Theta = \theta$, under EI and (4). Similarly

$$\frac{1}{J}\sum_{j=1}^{J}\sum_{m=0}^{n-1} \lambda''_{jm}(\theta)[Y_{jm} - P_{jm}(\theta)] \to 0.$$

Since $E[\frac{1}{J}\ell''_J(\theta)|\theta] = -\bar{I}_J(\theta)$, this implies that $\frac{1}{J}\ell''_J(\theta) + \bar{I}_J(\theta) \to 0$ in probability, given $\Theta = \theta$. Hence, using (4) again, we may rewrite (6) as

$$\frac{1}{J}\ell'_J(t) = o_p(1) - (t-\theta)[\bar{I}_J(\theta) - \frac{1}{2}(t-\theta)\rho_J M_\theta],$$

where ρ_J is a random quantity satisfying $|\rho_J| \le 1$, $o_p(1)$ denotes quantities tending to 0 in probability, given $\Theta = \theta$, and $\bar{I}_J(\theta)$ is bounded away from 0 by (5). Thus, for large J, $\frac{1}{J}\ell'_J(t)$ is approximately linear near θ, and with large probability is positive for some

$t_+ \in (\theta - \epsilon_0, \theta)$, negative for some $t_- \in (\theta, \theta + \epsilon_0)$, and by continuity equal to zero for some $\hat{\theta}_J \in (\theta - \epsilon_0, \theta + \epsilon_0$.) Hence, there is a sequence $\hat{\theta}_J$ of solutions to (2) with the property that for *any* $\epsilon_0 > 0$,

$$P[|\ \hat{\theta}_J - \theta\ | < \epsilon_0 \big| \Theta = \theta] \rightarrow 1,$$

as $J \rightarrow \infty$. Further details may be found in Serfling (1980, pp. 143–148). □

In general, we do not expect the roots $\hat{\theta}_J$ of (2) to be unique (e.g., Samejima, 1973). Moreover, among the multiple roots of (2), there is likely to be only one consistent root sequence: Foutz (1977) proves that if $\hat{\theta}_{1,J}$ and $\hat{\theta}_{2,J}$ are both consistent root sequences, then under LI, $P[\hat{\theta}_{1,J} = \hat{\theta}_{2,J}\ |\ \Theta = \theta] \rightarrow 1$ as $J \rightarrow \infty$. Thus, the situation, even under LI, is *opposite* that portrayed by Lord (1980, p. 59): rather than being optimistic that the roots should be eventually unique, one might be pessimistic that multiple roots continue to happen as $J \rightarrow \infty$, and only one of these roots for each J, brings us closer to the true θ.

This is not a practical problem, however. We shall see next that the standard practice of approximating a root of (2) by Newton's method, produces estimates that *are* consistent for θ under EI, even though the MLE and the Newton's method estimate of it were computed under the assumption LI. Thus, familiar numerical methods continue to be useful in estimating θ under EI.

Theorem 4.2. Suppose the assumptions of Theorem 4.1 hold, and let $\bar{\theta}_J$ be any sequence of consistent estimates for θ, given $\Theta = \theta$. Then the Newton's method improvement,

$$\theta_J^* = \bar{\theta}_J - \frac{\ell'_J(\bar{\theta}_J)}{\ell''_J(\bar{\theta}_J)}, \tag{7}$$

is also consistent for θ.

Proof: As in Lehmann (1983, p. 423) we may substitute a Taylor expansion of $\frac{1}{J}\ell'_J(\tilde{\theta}_J)$ about θ into (7) to obtain

$$\theta^*_J - \theta = -\frac{\frac{1}{J}\ell'_J(\theta)}{\frac{1}{J}\ell''_J(\tilde{\theta}_J)} + (\tilde{\theta}_J - \theta) \cdot o_p(1). \tag{8}$$

The second term on the right clearly tends to zero as $\tilde{\theta}_J \rightarrow \theta$. For the first term, a continuity argument shows that $-\frac{1}{J}\ell''_J(\tilde{\theta}_J) \approx \bar{I}_J(\theta) > 0$, and we know from (3) that $\frac{1}{J}\ell'_J(\theta) \rightarrow 0$. □

Clearly, the assertion of the theorem can be iterated to show that the result θ^*_J of, say, twenty Newton steps from $\tilde{\theta}_J$ would also be consistent. Such an estimator should be closer, in some sense, to the consistent roots found in Theorem 4.1. Newton's method requires an initial guess $\tilde{\theta}_J$; when LAD holds, $\tilde{\theta}_J = \bar{A}_J^{-1}(\bar{A}_J)$ is a natural choice, in view of Theorem 3.2.

5. Standard Error of the MLE

In the usual LI ability estimation theory, we expect that the sequence $\hat{\theta}_J$ will be *asymptotically normal* and *efficient*,

$$J^{\frac{1}{2}}(\hat{\theta}_J - \theta) \sim AN(0, 1/\bar{I}_J(\theta)), \tag{9}$$

as $J \rightarrow \infty$, where $\bar{I}_J(\theta)$ is the traditional test information function introduced in (6). A result like (9) identifying the standard error of $\hat{\theta}_J$ is needed to do statistical inference using $\hat{\theta}_J$—or indeed, merely to know how well to trust $\hat{\theta}_J$ as an estimator of θ for particular fixed J that arise in applications. However, (9) may fail in the essentially unidimensional case in two interesting ways: it may be that asymptotic normality holds but the asymptotic variance is no longer $\bar{I}_J(\theta)^{-1}$; or it may be that asymptotic normality fails completely.

When asymptotic normality does hold, we shall see that the deviation from the efficient variance is controlled by the quantity

$$C_J(\theta) = \frac{2}{J} \sum_{i=1}^{J} \sum_{j=1}^{i-1} \mathrm{Cov}(A_i, A_j \mid \Theta = \theta), \tag{10}$$

where the item scores A_j are constructed from the scoring scheme

$$a_{jm} \equiv \lambda'_{jm}(\theta), \ \forall\, j, m. \tag{11}$$

The scoring scheme (11) is a technical device that will be used throughout this section. The reader should not be mislead into thinking that (11) is a scoring scheme that could be applied to obtain a practical estimator as in Section 2 (to do so, we would already have to know θ!). Under EI and the bounds (4), we know that $\frac{1}{J}C_J(\theta) \to 0$, for all θ, but the behavior of $C_J(\theta)$ itself depends on the amount of local dependence in the item sequence $\underset{\sim}{X}$. Under LI of course, $C_J(\theta) \equiv 0$.

To see the effect of $C_J(\theta)$ on (9) under EI, we may deduce from (6) that

$$J^{\frac{1}{2}}(\hat{\theta}_J - \theta) \overset{\mathcal{D}}{\approx} \frac{J^{-\frac{1}{2}}\ell'_J(\theta)}{\bar{I}_J(\theta)}, \tag{12}$$

in the sense that the asymptotic distributions of the left and right hand sides are the same. If we can identify the asymptotic distribution of (a multiple of) $J^{-\frac{1}{2}}\ell'_J(\theta)$ then by (12), we will also be able to identify the asymptotic distribution and rate of convergence of $\hat{\theta}_J$. An indication of what is possible is provided by Theorem 5.1 below. Let us abbreviate

$$\sigma_J^2(\theta) = \mathrm{Var}(\bar{A}_J \mid \theta)$$

$$= \frac{1}{J^2} \sum_{j=1}^{J} \mathrm{Var}(A_j \mid \theta) + \frac{2}{J^2} \sum_{i=1}^{J} \sum_{j=1}^{i-1} \mathrm{Cov}(A_i, A_j \mid \theta), \tag{13}$$

for whatever scoring scheme $\{a_{jm}\}$ is currently under consideration. Under the scoring scheme (11), $\bar{A}_J - \bar{A}_J(\theta) \equiv \frac{1}{J}\ell'_J(\theta)$ and $\sigma_J^2(\theta) = \frac{1}{J}[\bar{I}_J(\theta) + C_J(\theta)]$.

Theorem 5.1. Suppose that the conditions of Theorem 4.1 hold for the item sequence $\underset{\sim}{X}$ and the latent trait Θ. Also, suppose that for some fixed θ, the scoring scheme (11) yields

$$\frac{1}{\sigma_J(\theta)}[\bar{A}_J - \bar{A}_J(\theta)] \sim AN(0,1). \tag{14}$$

Then, as $J \to \infty$,

(a) if $C_J(\theta) \to 0$ as $J \to \infty$,

$$J^{\frac{1}{2}}(\hat{\theta}_J - \theta) \sim AN\Big(0, \frac{1}{\bar{I}_J(\theta)}\Big);$$

(b) if $C_J(\theta)$ remains bounded for all J,

$$J^{\frac{1}{2}}(\hat{\theta}_J - \theta) \sim AN\Big(0, \frac{\bar{I}_J(\theta) + C_J(\theta)}{\bar{I}_J(\theta)^2}\Big);$$

(c) if $C_J(\theta)$ is unbounded and $R(J)$ is a function of J for which $R^2(J)C_J(\theta)/J$ remains bounded,

$$R(J)(\hat{\theta}_J - \theta) \sim AN\Big(0, \frac{R^2(J)C_J(\theta)}{J\bar{I}_J(\theta)^2}\Big).$$

Proof: From (12) and (14), only the asymptotic variance assertions need to be checked. For the scoring scheme (11), we have from (12), (13), and (14):

$$\begin{aligned} \mathrm{Var}(R(J)(\hat{\theta}_J - \theta)|\theta) &\approx \mathrm{Var}\Big(\frac{R(J)\ell'_J(\theta)}{J\bar{I}_J(\theta)} \mid \Theta = \theta\Big) \\ &= \frac{R^2(J)}{J\bar{I}_J(\theta)^2} \mathrm{Var}\Big\{\frac{1}{J}\sum_{j=1}^{J} A_i \mid \Theta = \theta\Big\} \\ &= \frac{R^2(J)}{J}\frac{\bar{I}_J(\theta) + C_J(\theta)}{\bar{I}_J(\theta)^2}. \end{aligned}$$

The assertions about the asymptotic variances in (a), (b), and (c) follow from this calculation by choosing $R(J)$ appropriately. □

Conditions (a) and (b) of the theorem correspond to the familiar case in which the rate of convergence of $\hat{\theta}_J$ to θ is $J^{-\frac{1}{2}}$. If $C_J(\theta) \to 0$ as $J \to \infty$, we get the usual asymptotically normal and efficient result (9) for $\hat{\theta}_J$. Otherwise, we get subefficiency or superefficiency depending on the sign of $C_J(\theta)$. The use of the terms efficient, subefficient, and superefficient to describe the asymptotic variance as being equal to, greater than, or less than $\bar{I}_J^{-1}(\theta)$ is suggestive here but perhaps misleading. In fact, $\bar{I}_J^{-1}(\theta)$ is the efficient variance only when $\ell_J(\theta)$ is the true log-likelihood function. Under EI, some other (unknown) log-likelihood function $L_J(\theta)$ applies, and examining the true efficiency of $\hat{\theta}_J$ would require access to the (unknown) function $E[-L_J''(\theta) \mid \theta]$.

Condition (c) corresponds to the rate of convergence of $\hat{\theta}_J$ to θ being *slower* than $J^{-\frac{1}{2}}$. This would happen, for example, if the inter-item covariances were generally positive and sufficiently large to force $C_J(\theta)$ to be unbounded. Formally, there is also the possibility that the convergence of $\hat{\theta}_J$ to θ could be *faster* than $J^{-\frac{1}{2}}$, but this would require that $\bar{I}_J(\theta) + C_J(\theta) \to 0$, that is, $C_J(\theta)$ negative for all large J. As we will argue next, this seems unlikely in many educational testing applications. Hence, this possibility was omitted from the theorem statement.

For reasonably homogeneous tests, one intuitively expects that items not independent given Θ would be positively correlated. This is certainly implicit in the factor-analytic tradition of test theory, (e.g., Anastasi, 1988, pp. 377 ff.). An example of the invocation of this principle in IRT research is the design of the simulation study in Drasgow and Parsons (1983). Indeed, it is quite reasonable to assume that if $\underset{\sim}{X}$ is essentially unidimensional with respect to the trait Θ, there are other traits $\Theta_2, \Theta_3, \ldots, \Theta_d$ such that LI holds with respect to the d-dimensional trait vector $(\Theta, \Theta_2, \Theta_3, \ldots, \Theta_d)$ (see, for example, Stout, 1989). If these traits are psychologically meaningful, it is also reasonable to assume that they will be ***associated*** (see Holland and Rosenbaum, 1986, for a definition), given $\Theta = \theta$. In the ordered-response case, a result of Jogdeo's (1978) can be used to argue that conditional on $\Theta = \theta$ alone, the inter-item covariances will be nonnegative (indeed, any ordered scoring scheme for $\underset{\sim}{X}$, given θ, will be associated). Thus, $C_J(\theta) > 0$ will generally be expected: the variance of $J^{\frac{1}{2}}(\hat{\theta}_J - \theta)$ will generally be higher than $1/\bar{I}_J(\theta)$.

Theorem 4.2 gave a practical way to approximate the estimator $\hat{\theta}_J$. The following corollary extends Theorem 5.1 to obtain asymptotic normality for this approximation.

Corollary 5.1. Suppose that the conditions of Theorem 5.1 hold, and that $\tilde{\theta}_J$ is any estimator with $R(J)(\tilde{\theta}_J - \theta)$ bounded in probability. The Newton's method approximation θ_J^* of Theorem 4.2 based on $\tilde{\theta}_J$ satisfies

$$R(J)(\theta_J^* - \theta) \sim \text{ AN } \left(0, \frac{R^2(J)}{J} \frac{\bar{I}_J(\theta) + C_J(\theta)}{\bar{I}_J(\theta)^2}\right).$$

Making appropriate choices of $R(J)$, we obtain the same three cases as in Theorem 5.1.

Proof: Using (11) and (13), we may rewrite (8) as

$$R(J)(\theta_J^* - \theta) \approx R(J) \frac{\bar{A}_J - \bar{A}_J(\theta)}{\sigma_J(\theta)} \frac{\{(\bar{I}_J(\theta) + C_J(\theta))/J\}^{1/2}}{\bar{I}_J(\theta)} + R(J)(\tilde{\theta}_J - \theta) o_p(1)$$

The result follows from (14), since $R(J)(\tilde{\theta}_J - \theta)$ is bounded in probability. □

By definition, $R(J)(\tilde{\theta}_J - \theta)$ is bounded in probability if $P[R(J) \mid \tilde{\theta}_J - \theta \mid \leq B \big| \Theta = \theta]$ can be made arbitrarily close to 1 as $J \to \infty$ by choosing B large enough. In Section 4 we suggested using $\tilde{\theta}_J = \bar{A}_J^{-1}(\bar{A}_J)$, for any convenient scoring scheme $\{a_{jm}\}$, as an initial guess for Newton's method under LAD. Routine calculation using Chebyshev's inequality shows that $R(J)(\tilde{\theta}_J - \theta)$ will then be bounded in probability as long as

$$\frac{R^2(J)}{J^2} \sum_{i=1}^{J} \sum_{j=1}^{i-1} \mathrm{Cov}(A_i, A_j \mid \Theta = \theta) \text{ is bounded,} \tag{15}$$

as $J \to \infty$. This represents a strengthening of EI since $R(J)$ is a fixed, increasing function of J for all θ. The assumption that (15) holds for scoring scheme (11) is also implicit in Theorem 5.1. Hereafter, we will say that *fast EI* holds if (15) holds for a fixed rate $R(J)$ and every bounded scoring scheme $\{a_{jm}\}$.

Theorem 5.1 also assumes that the raw score $\bar{A}_J$ is asymptotically normal in the sense of (20). Is this realistic? The following Central Limit Theorem (CLT) for dependent random variables, easily deduced from Theorem 2.2 of Dvoretzky (1972), sheds light on the qualitative side of this question.

Theorem 5.2. Suppose, for some fixed θ and some bounded scoring scheme $\{a_{jm}\}$:

(a) $J^2\sigma_J^2(\theta) \to \infty$;

(b) $\frac{1}{J\sigma_J(\theta)} \sum_{j=1}^{J} E[A_j - A_j(\theta) \mid \bar{A}_{j-1}, \theta] \to 0$; and

(c) $\frac{1}{J^2\sigma_J^2(\theta)} \sum_{j=1}^{J} \mathrm{Var}[A_j \mid \bar{A}_{j-1}, \theta] \to 1$,

as $J \to \infty$. Then, for this θ and $\{a_{jm}\}$,

$$\frac{1}{\sigma_J(\theta)}[\bar{A}_J - \bar{A}_J(\theta)] \sim AN(0,1). \tag{16}$$

The assumptions of Theorem 5.2 would be difficult to verify in practice, but this is somewhat offset by the fact that they are intuitively meaningful; thus, we can at least ask whether these assumptions are qualitatively appealing. Assumption (a) is merely a way of ensuring that most items contribute significantly to $\bar{A}_J$. It is difficult to imagine a useful item sequence or scoring scheme for which this would not be true. The conditioning in (b) is not only on a fixed value θ of Θ, but also on a fixed value of $\bar{A}_{j-1}$, for each j. If the conditioning on $\bar{A}_{j-1}$ were dropped, (b) would become an exact equality. Under EI, conditioning on $\Theta = \theta$ stabilizes $\bar{A}_{j-1}$ with high probability when j is large (Theorem 3.1), so we might expect that assumption (b) would hold for many EI item sequences. To gain some intuition about assumption (c), we may rewrite it as

$$\frac{\frac{1}{J}\sum_{j=1}^{J}\mathrm{Var}(A_j \mid \theta)}{\frac{1}{J}\sum_{j=1}^{J}\mathrm{Var}(A_j \mid \bar{A}_{j-1}, \theta)} + \frac{\frac{2}{J}\sum_{i=1}^{J}\sum_{j=1}^{i-1}\mathrm{Cov}(A_i, A_j \mid \theta)}{\frac{1}{J}\sum_{j=1}^{J}\mathrm{Var}(A_j \mid \bar{A}_{j-1}, \theta)} \to 1. \tag{17}$$

Hence, recalling that the A_j are bounded, (c) implies the fast EI condition,

$$\frac{2}{J}\sum_{i=1}^{J}\sum_{j=1}^{i-1}\mathrm{Cov}(A_i, A_j \mid \theta) \text{ is bounded, as } J \to \infty. \tag{18}$$

This condition is almost ubiquitous in general CLT's for dependent random variables (e.g., Bradley, 1985). Note that (18) precludes applying Theorem 5.2 in the situation of Theorem 5.1 (c); moreover, from (17) we can see that some additional balancing between the variance and covariance terms is needed for assumption (c) to hold. Example 5.2 below shows that Theorem 5.1 (c) can nevertheless occur.

There is another way in which the assumptions (b) and (c) are not entirely innocuous. EI and its strengthening (18) are second-order conditions (i.e., conditions that restrict only the expected values (given θ) of products of two item responses at a time). It is well-known that second-order conditions alone are not enough to guarantee (16). An example (without reference to latent traits) is constructed by Bradley (1989, Section 2) of a dichotomous sequence $\underset{\sim}{X}$ for which X_i and X_j are *independent* for every pair $i \neq j$ and yet the CLT fails. (The reader is referred to Bradley's paper for the rather complicated construction; also, see Bradley, 1985.) In light of the recent interest in Markov dependence among items (e.g., Jannarone, 1986; Spray & Ackerman, 1986), it is intriguing to observe that Bradley's example arises as a dichotomous scoring scheme for a Markov chain.

We conclude this section with two simpler examples illustrating the practical effects that item dependence can have on the standard error of $\hat{\theta}_J$. The examples are both variations of the paragraph comprehension example of Stout (1990; Example 2.3). Section 4.2 of Rosenbaum (1988) is also relevant. More complicated examples and/or examples in other realistic settings might also be constructed.

Example 5.1. Suppose $X_1, X_2, X_3, \ldots$ are binary item response variables, having the same response curve $P_j(\theta) = \theta$ (so the latent scale is the interval $(0,1)$ and $P[X_j = 1 \mid \theta] \equiv \theta$). Moreover, suppose that the items are arranged in successive groups of g_o items as

$$X_1, X_2, \ldots, X_{g_o};$$

$$X_{g_o+1}, X_{g_o+2}, \ldots, X_{2g_o};$$

$$etc.,$$

such that different groups of g_o items are independent of one another, given θ, and items within a single group are positively correlated, given θ. For simplicity, we will take

$$\mathrm{Corr}(X_i, X_j|\theta) = \begin{cases} c \text{ if } X_i \text{ and } X_j \text{ are in the same group,} \\ 0 \text{ if not,} \end{cases}$$

for some fixed $c \in (0,1]$. This is a naive model for a paragraph comprehension test in which several paragraphs are presented and g_o questions are asked for each paragraph. Here, θ represents a trait common to all the items, which we might wish to think of as reading comprehension; and the nonzero correlations are induced by nuisance traits, for example, specific knowledge about the subject matter of the paragraph at hand.

It is straightforward to verify that EI holds for this sequence of items; that is, for any bounded scoring scheme $\{a_{jm} : m = 0,1; j = 1,2,\ldots\}$ generating item scores $A_j = (1 - X_j)a_{j0} + X_j a_{j1}$,

$$\frac{1}{J^2}\sum_{i=1}^{J}\sum_{j=1}^{i-1}\mathrm{Cov}(A_i, A_j \mid \theta) \to 0.$$

Moreover, it can be verified (via Theorem 5.2 or by applying the usual CLT to the paragraph scores $G_k = \sum_{j=kg_o+1}^{(k+1)g_o} A_j$) that for any bounded scoring scheme $\{a_{jm}\}$ for which $J\sigma_J(\theta) \to \infty$,

$$\frac{1}{\sigma_J(\theta)}[\bar{A}_J - \bar{A}_J(\theta)] \sim AN(0,1),$$

given $\Theta = \theta$. Now, for the scoring scheme (11), the item scores are $A_j = (X_j - \theta)/\theta(1-\theta)$ so that $\mathrm{Cov}(A_i, A_j \mid \theta) = c/\theta(1-\theta)$ if A_i and A_j are in the same group, and 0 otherwise. Letting k_J be the greatest integer less than or equal to J/g_0, we see that

$$C_J(\theta) \approx \frac{2}{J}\sum_{k=1}^{k_J}\binom{g_0}{2}\frac{c}{\theta(1-\theta)} \approx \frac{c(g_0-1)}{\theta(1-\theta)},$$

and is bounded but nonzero as $J \to \infty$. Hence, using scoring scheme (11),

$$J^{\frac{1}{2}}(\hat{\theta}_J - \theta) \approx J^{\frac{1}{2}}\frac{\sigma_J(\theta)}{\bar{I}_J(\theta)}\frac{\bar{A}_J - \bar{A}_J(\theta)}{\sigma_J(\theta)},$$

and is asymptotically normal with mean 0 and variance

$$\frac{J\sigma_J^2(\theta)}{\bar{I}_J(\theta)^2} \approx \frac{\frac{1}{\theta(1-\theta)} + \frac{c(g_o-1)}{\theta(1-\theta)}}{[\frac{1}{\theta(1-\theta)}]^2} = \theta(1-\theta)[1 + c(g_o - 1)].$$

Note that the deviation from the efficient variance $\theta(1-\theta)$ is indeed due to dependence among the g_o items related to the same paragraph, and that $C_J(\theta) \approx c(g_o - 1)/\theta(1-\theta)$ appropriately characterizes this deviation. This illustrates part (b) of Theorem 5.1. □

The situation can be understood intuitively as follows: when items in a group are positively correlated given θ, a particular response to one item in the group is likely to be duplicated in responses to other items in the same group. Thus, a wrong response is likely to bias the θ-estimate downward more than is usual, and a right response is likely to bias the estimate upward, the biasing effect being magnified by the size of the group. This inflates the effect of noise inherent in the θ-estimation problem.

Example 5.2. Now let the sizes of the groups of mutually dependent items increase. We take dichotomous items $X_1, X_2, \ldots$ with identical ICC's $P_j(\theta) = \theta$ as before, but now group them as follows:

$$X_1, X_2, \ldots, X_{g(1)};$$

$$X_{g(1)+1}, X_{g(1)+2}, \ldots, X_{g(1)+g(2)};$$

etc.,

where $g(k)$ is a nondecreasing function of k. For specificity, we will take $g(k) \equiv k^{\frac{1}{4}}$. Once again, each group of $g(k)$ items is independent of the other groups, and for simplicity, we take $\mathrm{Corr}(X_i, X_j) \equiv c$ for X_i and X_j in the same group. We can verify that EI holds for this sequence of items, and apply Liapunov's Central Limit Theorem (Serfling, 1980, p.

30) to the paragraph scores

$$G_k = \sum_{\sum_1^k g(\ell)+1}^{\sum_1^{k+1} g(\ell)} A_j,$$

to conclude that

$$\frac{1}{\sigma_J(\theta)}[\bar{A}_J - \bar{A}_J(\theta)] \sim AN(0,1),$$

given $\Theta = \theta$, for any bounded scoring scheme $\{a_{jm}\}$. (Here, Theorem 5.2 does not apply at all, since we shall see below that $C_J(\theta) \to \infty$.)

As in the previous example

$$C_J(\theta) \approx \frac{2}{J}\sum_{k=1}^{k_J}\binom{g(k)}{2}\frac{c}{\theta(1-\theta)},$$

where k_J is chosen so that

$$\sum_{k=1}^{k_J} g(k) \leq J < \sum_{k=1}^{k_J+1} g(k),$$

that is, $k_J \approx (\frac{5}{4})^{\frac{4}{5}} J^{\frac{4}{5}}$ for $g(k) = k^{\frac{1}{4}}$. Thus, $C_J(\theta)$ grows like

$$\frac{\sum_{k=1}^{k_J} g(k)^2}{\sum_{k=1}^{k_J} g(k)} \approx \frac{5}{6}(\frac{5}{4})^{\frac{1}{5}} J^{\frac{1}{5}}.$$

(Incidentally, this also helps establish EI, since it shows that $C_J(\theta)/J \to 0$ as $J \to \infty$.)

Hence,

$$J^{\frac{2}{5}}(\hat{\theta}_J - \theta) \approx \frac{\{J^{-\frac{1}{5}}\bar{I}_J(\theta) + J^{-\frac{1}{5}}C_J(\theta)\}^{\frac{1}{2}}}{\bar{I}_J(\theta)} \cdot \frac{\bar{A}_J - \bar{A}_J(\theta)}{\sigma_J(\theta)},$$

which is asymptotically normal with mean zero and variance $J^{-\frac{1}{5}}C_J(\theta)/\bar{I}_J(\theta)^2 \approx (0.871)c\theta(1-\theta)$. Although the asymptotic variance appears lower than the efficient variance $\theta(1-\theta)$, the rate of convergence of $\hat{\theta}_J$ to θ is only $J^{-\frac{2}{5}}$, rather slower than the usual rate $J^{-\frac{1}{2}}$. □

In this example, the groups of dependent items become so large that the magnified effects of individual item responses have actually slowed the rate of convergence of $\hat{\theta}_J$ to θ. These magnified effects would be present in any θ-estimation method that ignored the nature of the inter-item dependencies. However, this need not be an argument against using estimation methods that assume local independence when this does not hold. The real lesson is that if one wants to continue to use a familiar estimator like $\hat{\theta}_J$ even though LI may fail, then one must be able to qualitatively justify an asymptotic distribution assumption like (14), and to quantitatively estimate $C_J(\theta)$ so that realistic standard errors of estimation can be calculated, etc. Note that in Example 5.2, $C_J(\theta)$ is unbounded as $J \to \infty$, but EI still holds; the unboundedness of $C_J(\theta)$ is responsible for the slower rate of convergence of $\hat{\theta}_J$ to θ. If $C_J(\theta)$ grows too fast as $J \to \infty$ then EI itself can also fail.

The quantity $C_J(\theta)$, or perhaps its average value over all θ's, should be viewed as an index of departure from local independence, locating collections of items—tests—along a continuum of unidimensional behavior from strictly locally independent unidimensional, $d_L = 1$, situations to dramatically non-unidimensional, $d_E > 1$, situations. This suggests the following model fit/trait estimation taxonomy, based upon the index $C_J(\theta)$ (contingent, of course, upon the qualitative acceptance of (14)):

I. $C_J(\theta) \approx 0$ *for all realistic* θ's. In this situation, ability estimation based on a $d_L = 1$ model could proceed as usual, using familiar standard errors such as $\bar{I}_J(\theta)^{-1/2}$. This situation covers both $d_L = 1$ settings as well as those essentially undimensional, $d_E = 1$, settings that only mildly violate LI.

II. $C_J(\theta) \neq 0$ *but moderate in size for all realistic* θ's. Here, Θ-estimation procedures based on LI could still be used but the conventional standard errors would have to be replaced by $(\bar{I}_J(\theta) + C_J(\theta))^{1/2}/\bar{I}_J(\theta)$. This would be the usual $d_E = 1$ setting.

III. $C_J(\theta) \neq 0$ *of substantial size for many* θ's. This would suggest that there is so much residual variability in the data after conditioning on Θ, that some genuinely multidimensional latent trait model may be needed.

Of course, the practical use of such a taxonomy rests on effective estimation of $C_J(\theta)$ itself. Work recently completed by Nandakumar and Stout (1989) aims at developing a practical index of EI for binary items, related to $C_J(\theta)$ but not adapted to the task of trait estimation. In particular, they investigate empirically the extent to which $d_E = 1$ holds or fails in the paragraph comprehension setting, as the number of items per paragraph increases.

Another approach to estimating $C_J(\theta)$ is suggested by the work of Gibbons, Bock, and Hedeker (1989). With the help of a computational device called the modified Clark algorithm, they are able to factor-analyze binary items, assumed to have normal ogive response curves, with correlated specific factors. $C_J(\theta)$ can then be estimated from the common factor loadings and specific factor correlations, at least when their one-factor solution leads to the same latent trait as identified in the definition of $d_E = 1$.

6. Application to the Dichotomous Case

In the binary (dichotomous) case, in which X_j takes the value 0 or 1 depending on the examinee's answer to the j^{th} item, the $d_L = 1$ likelihood is

$$P[\underset{\sim}{X}_J = \underset{\sim}{x}_J \mid \Theta = \theta] = \prod_{j=1}^{J} P_j(\theta)^{x_j}[1 - P_j(\theta)]^{1-x_j}, \tag{19}$$

with monotone item characteristic curves (ICC's) $P_j(\theta) = P[X_j = 1 \mid \Theta = \theta]$. Let us assume only that EI and LAD hold with respect to Θ. The definitions and theorems presented in Section 3 all specialize to the dichotomous setting, and in fact, most were introduced in this setting by Stout (1990). The MLE must solve the likelihood equation,

$$0 \equiv \ell'_J(\hat{\theta}_J) = \sum_{j=1}^{J} \lambda'_j(\hat{\theta}_J)[X_j - P_j(\hat{\theta}_J)], \tag{20}$$

where $\lambda_j(\theta) = \log P_j(\theta)/(1 - P_j(\theta))$ (the use of the log-odds-ratios λ_j is equivalent to using the log-category-probabilities λ_{j0} and λ_{j1} from Section 4, and avoids summation over the $n = 2$ response categories). As before, boundedness of $\lambda'_j(\theta)$ together with EI guarantees that $\frac{1}{J}\ell'_J(\theta)$ converges to zero, given $\Theta = \theta$. More precisely, we will assume that, for all θ, there exists an interval B_θ containing θ and a constant $M_\theta < \infty$, such that

$$|\lambda'''_j(t)| \leq M_\theta \ \forall\, t \in B_\theta, \ \forall\, j. \tag{21}$$

To complete a proof of consistency, we again need to bound the *test information function* as

$$\bar{I}_J(\theta) = \frac{1}{J}\sum_{j=1}^{J} \lambda'_j(\theta)P'_j(\theta) \geq \epsilon_\theta > 0, \tag{22}$$

as $J \to \infty$, and as in Proposition 4.1, LAD is a sufficient but not necessary condition to achieve this. Note that the information function in (22) is precisely the same one introduced in (5) for $n = 2$ response categories.

Theorem 6.1. Let $\underset{\sim}{X}$ be a dichotomous item sequence satisfying EI, (21), and (22). Then there exists a sequence $\{\hat{\theta}_J : J \geq J_0\}$ of roots of (20) such that

$$\lim_{J \to \infty} P[\,|\,\hat{\theta}_J - \theta\,| < \epsilon \mid \Theta = \theta] = 1,$$

for every $\epsilon > 0$.

Theorem 6.2. Suppose the assumptions of Theorem 6.1 hold, and let $\tilde{\theta}_J$ be any sequence of consistent estimates of θ, given $\Theta = \theta$. Then, the Newton's method improvement,

$$\theta_J^* = \tilde{\theta}_J - \frac{\ell'_J(\tilde{\theta}_J)}{\ell''_J(\tilde{\theta}_J)},$$

is also consistent for θ.

An obvious candidate for the initial guess in Theorem 6.2 is $\tilde{\theta}_J = \bar{P}_J^{-1}(\bar{X}_J)$. From (20) and the above results, we see again that the consistency and asymptotic distribution of $\hat{\theta}_J$ is tied up with the behavior of the centered weighted averages

$$\begin{aligned} \frac{1}{J}\ell'_J(\theta) &= \bar{A}_J - \bar{A}_J(\theta) \\ &= \frac{1}{J}\sum_{j=1}^{J} a_j[X_j - P_j(\theta)], \end{aligned} \tag{23}$$

with $a_j \equiv \lambda'_j(\theta)$, where again the dependence of a_j on θ does not matter since θ is fixed. Once again, let

$$\begin{aligned} \sigma_J^2(\theta) &= \mathrm{Var}(\bar{A}_J \mid \theta) \\ &= \frac{1}{J^2}\sum_{j=1}^{J} a_j^2 P_j(\theta)[1 - P_j(\theta)] + \frac{2}{J^2}\sum_{i=1}^{J}\sum_{J=1}^{i-1} a_i a_j \mathrm{Cov}(X_i, X_j \mid \Theta = \theta), \end{aligned}$$

and let $C_J(\theta) = (2/J)\sum_{i=1}^{J}\sum_{j=1}^{i-1} \lambda_i'(\theta)\lambda_j'(\theta)\mathrm{Cov}(X_i, X_j|\theta)$.

Theorem 6.3. Suppose that the assumptions of Theorem 6.1 hold for the item sequence $\underset{\sim}{X}$ and the latent trait Θ. Also suppose, given $\Theta = \theta$, that in (23),

$$\frac{1}{\sigma_J(\theta)}[\bar{A}_J - \bar{A}_J(\theta)] \sim AN(0,1). \tag{24}$$

Finally, suppose $R(J)$ is a function for which $R^2(J)C_J(\theta)/J$ remains bounded. Then,

$$R(J)(\hat{\theta}_J - \theta) \sim AN\left(0, \frac{R^2(J)}{J}\frac{\bar{I}_J(\theta) + C_J(\theta)}{\bar{I}_J(\theta)^2}\right).$$

Moreover, if $\tilde{\theta}_J$ is any estimator for which $R(J)(\tilde{\theta}_J - \theta)$ is bounded in probability, θ_J^* from Theorem 6.2 is also asymptotically normal with the same asymptotic variance.

Once (24) is deemed qualitatively acceptable, the asymptotic behavior of $\hat{\theta}_J$ is determined by $C_J(\theta)$. When $C_J(\theta)$ is near zero, we can expect the items to behave as though LI were true; when $C_J(\theta)$ is much larger, we should expect item behavior that can be effectively analyzed only with a multidimensional model.

7. Discussion

In assessing the shortcomings of the traditional local independence approach to item response modeling, Drasgow and Parsons (1983, p. 198) conclude, "it seems clear that researchers should be more concerned with the robustness of estimation techniques to minor violations of dimensionality assumptions than with the possibly neverending task of measuring all latent variables that underlie responses in a particular content domain." This call for the study of *structural robustness* in IRT is compelling: Although violations of strictly unidimensional latent structure can sometimes be explicitly modeled and exploited, many situations call for a unidimensional approach that is tolerant of minor violations of strict unidimensionality.

In this paper we have extended Stout's modeling notion of *essential independence*, EI, for binary items (Stout, 1987; 1990) to polytomous item sequences. Essential independence permits some dependence among items such as would be caused by minor violations of local independence, LI, due to nuisance trait multidimensionality, but still allows a single dominant latent trait to be identified. This type of mild interitem local dependence is arguably more realistic than unidimensional local independence models, $d_L = 1$, for many currently-used ability and achievement tests.

For items in which the expected amount of credit awarded increases with the latent trait, we have developed a theory of ability estimation under EI that closely parallels Stout (1990). As in Stout's dichotomous response theory, monotonicity need not be assumed for the individual items, but rather only for the test characteristic curve. Under this aggregate monotonicity condition, called *local asymptotic discrimination*, LAD, we have shown that the transformation $\bar{A}_J^{-1}(\bar{A}_J)$ of the raw test score is a consistent estimator of each examinee's θ as the test length J grows. A definition of *essential unidimensionality*, $d_E = 1$, was proposed based on EI and LAD holding with respect to a unidimensional trait Θ.

An alternative to scoring the items using an ad hoc scoring scheme $\{a_{jm}\}$ (which leads to the test scores $\bar{A}_J$ above) is to ignore the local dependence among the items and employ a well-known LI-based estimation procedure. Since it is common to use an LI model even when LI is believed to be only approximately true, the behavior of such a procedure in the more realistic EI setting is an important issue, as Drasgow and Parsons attest to above. Maximum likelihood estimation of θ was examined in this light.

The MLE $\hat{\theta}_J$ based on a unidimensional LI model was shown to be consistent for each examinee's θ as the test length J grows, when only EI and not LI holds. In this sense $\hat{\theta}_J$ is robust as an estimator of θ under this realistic structural violation of LI. When an estimator such as $\hat{\theta}_J$ is found to be consistent, its precision as an estimator is usually judged by the theoretical asymptotic distribution of $J^{1/2}(\hat{\theta}_J - \theta)$. Under LI, we expect this to be asymptotically normal with mean zero and variance $1/\bar{I}_J(\theta)$ as $J \to \infty$, where $\bar{I}_J(\theta)$ is the test information function. When $\hat{\theta}_J$ is based on an LI model but only EI holds, this asymptotic distribution may fail in various ways: the rate $J^{1/2}$ may be preserved but the variance may be inflated by an essentially constant amount; the rate $J^{1/2}$ may fail; and finally, it is conceivable that asymptotic normality itself fails, with any rate of convergence. Hence, the robustness of consistency for the MLE does not extend to a robustness of asymptotic distribution, under EI violations of LI.

Conditions for asymptotic normality of $\hat{\theta}_J$ involve higher product-moment assumptions that do not admit easy rigorous checks from the data. Hence, asymptotic normality itself is usually a qualitative issue that must be decided by the practitioner in each application. If asymptotic normality is qualitatively acceptable, the correct variance can be calculated with the help of the expression

$$C_J(\theta) = \frac{2}{J} \sum_{i=1}^{J} \sum_{j=1}^{i-1} \mathrm{Cov}(A_i, A_j \mid \theta),$$

where the A_j's score each response category according to the derivative of the log-category-probability: $A_j = \sum_{m=0}^{n-1} \lambda'_{jm}(\theta) Y_{jm}$. In principle, $C_J(\theta)$ could be positive or negative; however, in many educational testing settings, we expect it to be positive. Under EI, $C_J(\theta)/J \to 0$, but $C_J(\theta)$ itself need not tend to zero.

For fixed J, the quantity $C_J(\theta)$ should be viewed as an index of local item dependence along a continuum that connects strictly $d_L = 1$ unidimensional models with strictly $d_E > 1$ multidimensional models. Such a continuum has also been suggested by Drasgow and Parsons (1983). The $d_E = 1$ unidimensional models, which are the focus of this paper, form the middle of this continuum. The nearer $C_J(\theta)$ to zero, the more we can expect latent trait estimation to behave as though LI were true. The larger $C_J(\theta)$, the more we should expect item behavior that can be effectively analyzed only with an explicitly multidimensional latent trait model. Thus, if $C_J(\theta)$ could be effectively estimated in practice, we would be able to use it to predict the behavior of $\hat{\theta}_J$. Various ideas for doing this are provided by Wainer and Wright (1980), Gibbons, Bock, and Hedeker (1989), and Nandakumar and Stout (1989). For this reason, the non-robustness of distribution of $\hat{\theta}_J$ need not be defeating.

The principal assumptions needed to establish consistency of the LI-based MLE were EI and that the information function $\bar{I}_J(\theta)$ (calculated as though LI were true) be bounded away from 0 and ∞. Indeed, a hierarchy of identifiability conditions for estimating θ can be developed, starting with cumulative RCC monotonicity M (i.e., ordered-response items), moving through test characteristic curve monotonicity LAD, to the bounding of $\bar{I}_J(\theta)$ away from 0. Each of these conditions in some sense implies the next, and all allow various forms of unidimensional latent trait estimation. This hierarchy illustrates the transition from highly interpretable but very restrictive conditions, such as M, to less restrictive conditions that do not admit easy psychometric interpretation, such as the bounding conditions on $\bar{I}_J(\theta)$.

Essential independence plays a central role in the convergence of $\hat{\theta}_J$ to θ because it guarantees the stability of certain weighted averages of item scores that appear in the LI-based log-likelihood. Therefore, we might expect that under EI and suitable regularity conditions, other estimators that depend on the stability of the LI-based log-likelihood would also be consistent estimators of θ. Indeed, a trivial modification of the proof of Theorem 4.1 shows that the posterior mode, which maximizes the posterior density

$$f_J(\theta \mid \underset{\sim}{X}_J) = \frac{P[\underset{\sim}{X}_J = \underset{\sim}{x}_J \mid \theta] f(\theta)}{P[\underset{\sim}{X}_J = \underset{\sim}{x}_J]},$$

is consistent for θ under the conditions of that theorem and a mild nondegeneracy condition on the density $f(\theta)$ of Θ in the examinee population. The posterior mode has been considered by Samejima (1969) and by Lord (1986), for example. A different set of regularity conditions from those employed in Theorem 4.1, which are equally plausible in applications, can be used to obtain consistency of the posterior mean,

$$E[\Theta \mid \underset{\sim}{X}_J] = \int \theta f_J(\theta \mid \underset{\sim}{X}_J) d\theta.$$

Essential independence is used here to ignore the part of the integral away from the value θ_0 that generated the data $\underset{\sim}{X}_J$; see, for example, the proof of equation (5) in Walker (1969). The regularity conditions needed generalize Walker's conditions, and incidentally provide another proof of consistency of the MLE. The posterior mean has been considered by, for example, Bock and Mislevy (1982) as well as earlier by Samejima (1969).

Essential independence is thus seen to be a minimal condition under which strictly $d_L = 1$ trait estimation procedures may be expected to work when applied to mildly multidimensional data. Our examination of essential independence in the polytomous item response setting shows that this condition is not an artifact of the simple structure of dichotomously-scored tests, but a general condition that can be fruitfully applied to standardized tests of all sorts. Moreover, we have shown that a rigorous approach to the structural robustness analysis advocated by Drasgow and Parsons (1983) is possible. Locally independent latent trait models can, and should, continue to be used to develop estimation and decision procedures in IRT, if for no other reason than their analytic simplicity. However, before LI-based procedures are applied on-line, they should be thoroughly examined under the more realistic assumption of essential independence.

References

Anastasi, A. (1988). *Psychological testing* (sixth edition). New York: Macmillan Publishing Company.

Ash, R. B. (1972). *Real analysis and probability.* New York: Academic Press.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In Lord, F. M. & M. R. Novick, *Statistical theories of mental test scores.* Reading, Massachusettes: Addison-Wesley.

Bock, R. D. (1972). Estimating item parameters and latent ability when the responses are scored in two or more nominal categories. *Psychometrika, 37,* 29–51.

Bock, R. D., & R. J. Mislevy (1982). Adaptive EAP estimation of ability in a microcomputer environment. *Applied Psychological Measurement, 6,* 431–444.

Bradley, R. C. (1985). Counterexamples to the central limit theorem under strong mixing conditions. In P. Revesz (Ed.), *Colloquia mathematica societitas Janos Bolyai 36. Limit theorems in probability and statistics, Veszprem (Hungary), 1982.* Volume I (pp. 153–172). Amsterdam: North-Holland.

Bradley, R. C. (1989). A stationary, pairwise independent, absolutely regular sequence for which the central limit theorem fails. *Probability Theory and Related Fields, 81,* 1–10.

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

Dvoretzky, A. (1972). Asymptotic normality for sums of dependent random variables. *Proceedings of the sixth Berkeley symposium on mathematical probability and statistics,* Volume II (pp. 513–535).

Foutz, R. V. (1977). On the unique consistent solution to the likelihood equations. *Journal of the American Statistical Association, 72,* 147–148.

Gibbons, R. D., R. D. Bock, & D. R. Hedeker (1989). *Conditional dependence.* Final Research Report, Office of Naval Research and Illinois State Psychiatric Institute. Chicago, Illinois: University of Illinois at Chicago.

Holland, P. W., & P. R. Rosenbaum (1986). Conditional association and unidimensionality in monotone latent trait models. *Annals of Statistics, 14*, 1523–1543.

Jannarone, R. J. (1986). Conjunctive item response theory kernels. *Psychometrika, 51*, 357–373.

Jogdeo, K. (1978). On a probability bound of Marshall and Olkin. *Annals of Statistics, 6*, 232–234.

Junker, B. W. (1988). *Statistical aspects of a new latent trait model.* Ph.D. thesis, Department of Statistics. Champaign, Illinois: University of Illinois at Urbana-Champaign.

Lehmann, E. L. (1983). *Theory of point estimation.* New York: Wiley.

Levine, M. V. (1989). *Ability distributions, pattern probabilities, and quasidensities.* Final Research Report, Office of Naval Research and Model-Based Measurement Laboratory. Champaign, Illinois: University of Illinois at Urbana-Champaign.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Lord, F. M. (1986). Maximum likelihood and Bayesian parameter estimation in item response theory. *Journal of Educational Measurement, 23*, 157–162.

Lord, F. M. & M. R. Novick (1968). *Statistical theories of mental test scores.* Reading, Massachusettes: Addison-Wesley.

Nandakumar, R. & W. F. Stout (1989). Traditional dimensionality vs. essential dimensionality. Submitted to *Journal of Educational Measurement.*

Rosenbaum, P. R. (1988). Item bundles. *Psychometrika, 53*, 349–359.

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores. Psychometrika* Monograph Supplement to Volume 24, Monograph Number 17.

Samejima, F. (1972). *A general model for free-response data. Psychometrika* Monograph

Supplement to Volume 37, Monograph Number 18.

Samejima, F. (1973). A comment on Birnbaum's three parameter logistic model in the latent trait theory. *Psychometrika, 38,* 221–233.

Serfling, R. J. (1980). *Approximation theorems in mathematical statistics.* New York: Wiley.

Spray, J. A., & T. A. Ackerman (1987). *The effect of item response dependency on trait or ability dimensionality.* ACT research report series #87–10, Iowa City, Iowa.

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

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

Thissen, D., & L. Steinberg (1986). A taxonomy of item response models. *Psychometrika, 51,* 567–577.

Tsutakawa, R. K. & M. J. Soltys (1988). Approximation for Bayesian ability estimation. *Journal of Educational Statistics, 13,* 117–130.

Wainer, H. & B. D. Wright (1980). Robust estimation of ability in the Rasch model. *Psychometrika, 45,* 373–391.

Walker, A. M. (1969). On the asymptotic behaviour of posterior distributions. *Journal of the Royal Statistical Society, Series B, 31,* 80–88.

Distribution List

Dr. Terry Ackerman
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. James Algina
1403 Norman Hall
University of Florida
Gainesville, FL 32605

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

Dr. Ronald Armstrong
Rutgers University
Graduate School of Management
Newark, NJ 07102

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Laura L. Barnes
College of Education
University of Toledo
2801 W. Bancroft Street
Toledo, OH 43606

Dr. William M. Bart
University of Minnesota
Dept. of Educ. Psychology
330 Burton Hall
178 Pillsbury Dr., S.E.
Minneapolis, MN 55455

Dr. Isaac Bejar
Law School Admissions
Services
P.O. Box 40
Newtown, PA 18940-0040

Dr. Ira Bernstein
Department of Psychology
University of Texas
P.O. Box 19528
Arlington, TX 76019-0528

Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Ramat Aviv 69978
ISRAEL

Dr. Arthur S. Blaiwes
Code N712
Naval Training Systems Center
Orlando, FL 32813-7100

Dr. Bruce Bloxom
Defense Manpower Data Center
99 Pacific St.
Suite 155A
Monterey, CA 93943-3231

Cdt. Arnold Bohrer
Sectie Psychologisch Onderzoek
Rekruterings-En Selectiecentrum
Kwartier Koningen Astrid
Bruijnstraat
1120 Brussels, BELGIUM

Dr. Robert Breaux
Code 281
Naval Training Systems Center
Orlando, FL 32826-3224

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

Dr. Gregory Candell
CTB/McGraw-Hill
2500 Garden Road
Monterey, CA 93940

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

Dr. John M. Carroll
IBM Watson Research Center
User Interface Institute
P.O. Box 704
Yorktown Heights, NY 10598

Dr. Robert M. Carroll
Chief of Naval Operations
OP-01B2
Washington, DC 20350

Dr. Raymond E. Christal
UES LAMP Science Advisor
AFHRL/MOEL
Brooks AFB, TX 78235

Mr. Hua Hua Chung
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Norman Cliff
Department of Psychology
Univ. of So. California
Los Angeles, CA 90089-1061

Director, Manpower Program
Center for Naval Analyses
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Hans F. Crombag
Faculty of Law
University of Limburg
P.O. Box 616
Maastricht
The NETHERLANDS 6200 MD

Ms. Carolyn R. Crone
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Timothy Davey
American College Testing Program
P.O. Box 168
Iowa City, IA 52243

Dr. C. M. Dayton
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Ralph J. DeAyala
Measurement, Statistics,
and Evaluation
Benjamin Bldg., Rm. 4112
University of Maryland
College Park, MD 20742

Dr. Lou DiBello
CERL
University of Illinois
103 South Mathews Avenue
Urbana, IL 61801

Dr. Dattprasad Divgi
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Mr. Hei-Ki Dong
Bell Communications Research
Room PYA-1K207
P.O. Box 1320
Piscataway, NJ 08855-1320

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

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
(2 Copies)

Dr. Stephen Dunbar
224B Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. James A. Earles
Air Force Human Resources Lab
Brooks AFB, TX 78235

Dr. Susan Embretson
University of Kansas
Psychology Department
426 Fraser
Lawrence, KS 66045

Dr. George Englehard, Jr.
Division of Educational Studies
Emory University
210 Fishburne Bldg.
Atlanta, GA 30322

ERIC Facility-Acquisitions
2440 Research Blvd, Suite 550
Rockville, MD 20850-3238

Dr. Benjamin A. Fairbank
Operational Technologies Corp.
5825 Callaghan, Suite 225
San Antonio, TX 78228

Dr. Marshall J. Farr, Consultant
Cognitive & Instructional Sciences
2520 North Vernon Street
Arlington, VA 22207

Dr. P-A. Federico
Code 51
NPRDC
San Diego, CA 92152-6800

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

Dr. Richard L. Ferguson
American College Testing
P.O. Box 168
Iowa City, IA 52243

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

Dr. Myron Fischl
U.S. Army Headquarters
DAPE-MRR
The Pentagon
Washington, DC 20310-0300

Prof. Donald Fitzgerald
University of New England
Department of Psychology
Armidale, New South Wales 2351
AUSTRALIA

Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Alfred R. Fregly
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Robert D. Gibbons
Illinois State Psychiatric Inst.
Rm 529W
1601 W. Taylor Street
Chicago, IL 60612

Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01003

Dr. Drew Gitomer
Educational Testing Service
Princeton, NJ 08541

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

Dr. Sherrie Gott
AFHRL/MOMJ
Brooks AFB, TX 78235-5601

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

Michael Habon
DORNIER GMBH
P.O. Box 1420
D-7990 Friedrichshafen 1
WEST GERMANY

Prof. Edward Haertel
School of Education
Stanford University
Stanford, CA 94305

Dr. Ronald K. Hambleton
University of Massachusetts
Laboratory of Psychometric
and Evaluative Research
Hills South, Room 152
Amherst, MA 01003

Dr. Delwyn Harnisch
University of Illinois
51 Gerty Drive
Champaign, IL 61820

Dr. Grant Henning
Senior Research Scientist
Division of Measurement
Research and Services
Educational Testing Service
Princeton, NJ 08541

Ms. Rebecca Hetter
Navy Personnel R&D Center
Code 63
San Diego, CA 92152-6800

Dr. Thomas M. Hirsch
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Paul W. Holland
Educational Testing Service, 21-T
Rosedale Road
Princeton, NJ 08541

Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 92010

Ms. Julia S. Hough
Cambridge University Press
40 West 20th Street
New York, NY 10011

Dr. William Howell
Chief Scientist
AFHRL/CA
Brooks AFB, TX 78235-5601

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

Dr. Steven Hunka
3-104 Educ. N.
University of Alberta
Edmonton, Alberta
CANADA T6G 2G5

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

Dr. Robert Jannarone
Elec. and Computer Eng. Dept.
University of South Carolina
Columbia, SC 29208

Dr. Kumar Joag-dev
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright Street
Champaign, IL 61820

Dr. Douglas H. Jones
1280 Woodfern Court
Toms River, NJ 08753

Dr. Brian Junker
Carnegie-Mellon University
Department of Statistics
Schenley Park
Pittsburgh, PA 15213

Dr. Michael Kaplan
Office of Basic Research
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Milton S. Katz
European Science Coordination
Office
U.S. Army Research Institute
Box 65
FPO New York 09510-1500

Prof. John A. Keats
Department of Psychology
University of Newcastle
N.S.W. 2308
AUSTRALIA

Dr. Jwa-keun Kim
Department of Psychology
Middle Tennessee State
University
P.O. Box 522
Murfreesboro, TN 37132

Mr. Soon-Hoon Kim
Computer-based Education
Research Laboratory
University of Illinois
Urbana, IL 61801

Dr. G. Gage Kingsbury
Portland Public Schools
Research and Evaluation Department
501 North Dixon Street
P. O. Box 3107
Portland, OR 97209-3107

Dr. William Koch
Box 7246, Meas. and Eval. Ctr.
University of Texas-Austin
Austin, TX 78703

Dr. Richard J. Koubek
Department of Biomedical
& Human Factors
139 Engineering & Math Bldg.
Wright State University
Dayton, OH 45435

Dr. Leonard Kroeker
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Dr. Jerry Lehnus
Defense Manpower Data Center
Suite 400
1600 Wilson Blvd
Rosslyn, VA 22209

Dr. Thomas Leonard
University of Wisconsin
Department of Statistics
1210 West Dayton Street
Madison, WI 53705

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

Dr. Charles Lewis
Educational Testing Service
Princeton, NJ 08541-0001

Mr. Rodney Lim
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert L. Linn
Campus Box 249
University of Colorado
Boulder, CO 80309-0249

Dr. Robert Lockman
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

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

Dr. Richard Luecht
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. George B. Macready
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

Dr. Clessen J. Martin
Office of Chief of Naval
Operations (OP 13 F)
Navy Annex, Room 2832
Washington, DC 20350

Dr. James R. McBride
HumRRO
6430 Elmhurst Drive
San Diego, CA 92120

Dr. Clarence C. McCormick
HQ, USMEPCOM/MEPCT
2500 Green Bay Road
North Chicago, IL 60064

Mr. Christopher McCusker
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert McKinley
Educational Testing Service
Princeton, NJ 08541

Mr. Alan Mead
c/o Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Timothy Miller
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Dr. William Montague
NPRDC Code 13
San Diego, CA 92152-6800

Ms. Kathleen Moreno
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Headquarters Marine Corps
Code MPI-20
Washington, DC 20380

Dr. Ratna Nandakumar
Educational Studies
Willard Hall, Room 213E
University of Delaware
Newark, DE 19716

Library, NPRDC
Code P201L
San Diego, CA 92152-6800

Librarian
Naval Center for Applied Research
in Artificial Intelligence
Naval Research Laboratory
Code 5510
Washington, DC 20375-5000

Dr. Harold F. O'Neil, Jr.
School of Education - WPH 801
Department of Educational
Psychology & Technology
University of Southern California
Los Angeles, CA 90089-0031

Dr. James B. Olsen
WICAT Systems
1875 South State Street
Orem, UT 84058

Office of Naval Research,
Code 1142CS
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Dr. Judith Orasanu
Basic Research Office
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311

Dr. Peter J. Pashley
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

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

Dr. James Paulson
Department of Psychology
Portland State University
P.O. Box 751
Portland, OR 97207

Dept. of Administrative Sciences
Code 54
Naval Postgraduate School
Monterey, CA 93943-5026

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

Dr. Malcolm Ree
AFHRL/MOA
Brooks AFB, TX 78235

Mr. Steve Reise
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

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

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

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
310B Austin Peay Bldg.
Knoxville, TN 37916-0900

Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152-6800

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

Dr. Mary Schratz
4100 Parkside
Carlsbad, CA 92008

Dr. Dan Segall
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Robin Shealy
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

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

Dr. Randall Shumaker
Naval Research Laboratory
Code 5510
4555 Overlook Avenue, S.W.
Washington, DC 20375-5000

Dr. Richard E. Snow
School of Education
Stanford University
Stanford, CA 94305

Dr. Richard C. Sorensen
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Judy Spray
ACT
P.O. Box 168
Iowa City, IA 52243

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Dr. Peter Stoloff
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. William Stout
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

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

Mr. Brad Sympson
Navy Personnel R&D Center
Code-62
San Diego, CA 92152-6800

Dr. John Tangney
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Kikumi Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Maurice Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

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

Mr. Thomas J. Thomas
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Mr. Gary Thomasson
University of Illinois
Educational Psychology
Champaign, IL 61820

Dr. Robert Tsutakawa
University of Missouri
Department of Statistics
222 Math. Sciences Bldg.
Columbia, MO 65211

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

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 440
St. Paul, MN 55114

Dr. Frank L. Vicino
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

Dr. Michael T. Waller
University of Wisconsin-Milwaukee
Educational Psychology Department
Box 413
Milwaukee, WI 53201

Dr. Ming-Mei Wang
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Thomas A. Warm
FAA Academy AAC934D
P.O. Box 25082
Oklahoma City, OK 73125

Dr. Brian Waters
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. Ronald A. Weitzman
Box 146
Carmel, CA 93921

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Douglas Wetzel
Code 51
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Rand R. Wilcox
University of Southern
California
Department of Psychology
Los Angeles, CA 90089-1061

German Military Representative
ATTN: Wolfgang Wildgrube
Streitkraefteamt
D-5300 Bonn 2
4000 Brandywine Street, NW
Washington, DC 20016

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

Dr. Hilda Wing
Federal Aviation Administration
800 Independence Ave, SW
Washington, DC 20591

Mr. John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering
Cancer Center
1275 York Avenue
New York, NY 10021

Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
Code 51
San Diego, CA 92152-6800

Dr. Kentaro Yamamoto
02-T
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

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

Dr. Joseph L. Young
National Science Foundation
Room 320
1800 G Street, N.W.
Washington, DC 20550

Mr. Anthony R. Zara
National Council of State
Boards of Nursing, Inc.
625 North Michigan Avenue
Suite 1544
Chicago, IL 60611