UNBIASED ESTIMATORS OF ABILITY PARAMETERS,
OF THEIR VARIANCE, AND OF THEIR PARALLEL-FORMS RELIABILITY

Frederic M. Lord

This research was sponsored in part by the Personnel and Training Research Programs Psychological Sciences Division Office of Naval Research, under Contract No. N00014-80-C-0402

Contract Authority Identification Number NR No. 150-453

Frederic M. Lord, Principal Investigator

Educational Testing Service Princeton, New Jersey

November 1981

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for public release; distribution unlimited.
UNBIASED ESTIMATORS OF ABILITY PARAMETERS,
OF THEIR VARIANCE, AND OF THEIR
PARALLEL-FORMS RELIABILITY

Frederic M. Lord

This research was sponsored in part by the Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research, under Contract No. N00014-80-C-0402
Contract Authority Identification Number NR No. 150-453
Frederic M. Lord, Principal Investigator

Educational Testing Service
Princeton, New Jersey

November 1981

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for public release; distribution unlimited.
<table>
<thead>
<tr>
<th>Report Number</th>
<th>AD-A108874</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Unbiased Estimators of Ability Parameters, of Their Variance, and of Their Parallel-Forms Reliability</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Frederic M. Lord</td>
</tr>
<tr>
<td>Performing Organization Name and Address</td>
<td>Educational Testing Service Princeton, NJ 08541</td>
</tr>
<tr>
<td>Controlling Office Name and Address</td>
<td>Personnel and Training Research Programs Office of Naval Research (Code 458)</td>
</tr>
<tr>
<td>Type of Report &amp; Period Covered</td>
<td>Technical Report</td>
</tr>
<tr>
<td>Research Report Number</td>
<td>Research Report 81-50</td>
</tr>
<tr>
<td>Contract or Grant Number(s)</td>
<td>N00014-80-C-0402</td>
</tr>
<tr>
<td>Program Element, Project, Task Area &amp; Work Unit Numbers</td>
<td>NR 150-453</td>
</tr>
<tr>
<td>Report Date</td>
<td>November 1981</td>
</tr>
<tr>
<td>Number of Pages</td>
<td>29</td>
</tr>
<tr>
<td>Distribution Statement (of this report)</td>
<td>Approved for public release; distribution unlimited.</td>
</tr>
<tr>
<td>Distribution Statement (of the abstract entered in Block 20, if different from Report)</td>
<td></td>
</tr>
<tr>
<td>Supplementary Notes</td>
<td></td>
</tr>
<tr>
<td>Key Words</td>
<td>Bias (statistical), Estimation, Item Response Theory, Ability, True Score, Reliability, Maximum Likelihood, Mental Test Theory, Unbiased Estimate, Standard Error, Asymptotics</td>
</tr>
<tr>
<td>Abstract (Continue on reverse side if necessary and identify by block number)</td>
<td>Given known item parameters, unbiased estimators are derived 1) for an examinee's ability parameter $\theta$ and for his proportion-correct true score $\xi$, 2) for $\sigma^2_0$ the variance of $\theta$ across examinees in the group tested, also for $\sigma^2_\xi$, and 3) for the parallel-forms reliability of the observed test score, the maximum likelihood estimator $\hat{\theta}$.</td>
</tr>
</tbody>
</table>
Unbiased Estimators

Abstract

Given known item parameters, unbiased estimators are derived
1) for an examinee's ability parameter $\theta$ and for his proportion-
correct true score $\zeta$, 2) for $s_\theta^2$ the variance of $\theta$ across examinees
in the group tested, also for $s_\zeta^2$, and 3) for the parallel-forms
reliability of the observed test score, the maximum likelihood estimator,
$\hat{\theta}$. 
Unbiased Estimators of Ability Parameters, of Their Variance, and of Their Parallel-Forms Reliability

This paper is primarily concerned with determining the statistical bias in the maximum likelihood estimate \( \theta \) of the examinee ability parameter \( \theta \) in item response theory (IRT) [Lord, 1980]; also of certain functions of such parameters. We will deal only with unidimensional tests composed of dichotomously scored items. We assume the item response function is three-parameter logistic (2).

Available results for the sampling variance of \( \theta \) are currently limited to the case where the item parameters are known; the present derivations are limited to this case also. This limitation is tolerable in situations where the item parameters are predetermined, as in item banking and tailored testing.

In the absence of a prior distribution for \( \theta \), it is well known that examinees with perfect scores have \( \hat{\theta} = \infty \); also that examinees who perform near or below the chance level on multiple-choice items may be given large negative values of \( \hat{\theta} \). This (correctly) suggests that \( \hat{\theta} \) is positively biased for high-ability examinees and negatively biased for low-ability examinees. Will a correction of \( \hat{\theta} \) for bias be helpful in such cases?

*This work was supported in part by contract N00014-80-C-0402, project designation NR 150-453 between the Office of Naval Research and Educational Testing Service. Reproduction in whole or in part is permitted for any purpose of the United States Government.
Unbiased Estimators

It is also 'well known' that for any ordinary group of examinees, the variance \((s^2_\theta)\) of \(\hat{\theta}\) across examinees is larger than the variance \((s^2_0)\) of the true \(\theta\). The ratio \(s^2_\theta/s^2_0\) is closely related to the classical-test-theory reliability of \(\theta\) considered as the examinee's test score. Thus it is not enough for us to know that \(s^2_\theta > s^2_0\) as the number \(n\) of test items becomes large; we need to know how the relation of \(s^2_\theta\) to \(s^2_0\) varies as a function of \(n\). We also need a better estimate of \(s^2_\theta\) than its maximum-likelihood estimator \(s^2_\theta\).

These objectives can be achieved by correcting \(s^2_\theta\) for bias.

The methods used to derive formulas for correction for bias are presented here in detail for at least two reasons: 1) experience with similar derivations has shown that it is easy to reach erroneous results if details are not spelled out. 2) The general methods used here are easily transferred to solve other problems, such as a) correction of item parameters for bias, b) obtaining higher-order approximations to the sampling variance of \(\hat{\theta}\).

1. Statistical Bias in \(\hat{\theta}\) and \(\hat{\xi}\)

The method used here to find the bias of \(\hat{\theta}\) is adapted from the 'adjusted order of magnitude' procedure detailed by Shenton and Bowman (1977). They assume their data to be a sample from a population divided into a denumerable number of subsets. For them, the population proportion of observations in a given subset is a known function of the parameter \(\theta\) whose value they wish to estimate. Their sample estimate of \(\theta\) is therefore a function of observed sample proportions in the
various subsets. Since our data do not readily fit this picture, we cannot use their final published formulas but must instead derive our own.

Throughout Section 1, we deal with a single fixed examinee whose ability $\theta$ is the parameter to be estimated. All item parameters are assumed known.

1.1 Preliminaries

The maximum likelihood estimate $\hat{\theta}$ is obtained by solving the likelihood equation

$$\sum_{i=1}^{n} (u_i - \hat{P}_i)\hat{P}_i' / \hat{P}_i \hat{Q}_i = 0$$

(1)

where $u_i = 0$ or 1 is the examinee's response to item $i$ ($i = 1, 2, \ldots, n$), $P_i \equiv P_i(\theta)$ is the response function for item $i$, $Q_i \equiv 1 - P_i$, $P_i'$ is the derivative of $P_i$ with respect to $\theta$, and a caret indicates that the function is to be evaluated at $\hat{\theta}$. We deal with the case where $P_i$ is the three-parameter logistic function

$$P_i \equiv c_i + \frac{1 - c_i}{1 + e^{-A_i(\theta - b_i)}}$$

(2)

where $A_i$, $b_i$, and $c_i$ are item parameters describing item $i$. 
We will assume

1. $\theta$ is a bounded variable,
2. the item parameters $a_i$ and $b_i$ are bounded,
3. $c_i$ is bounded away from 1,
4. as $n$ becomes large, the statistical characteristics of the test stabilize.

Rather than trying to define this last assumption formally, the reader may substitute the more restrictive assumption usually made in mental test theory: that a test is lengthened by adding strictly parallel forms.

With these assumptions, the conditions of Bradley and Gart (1962) are satisfied. It follows from their theorems that $\hat{\theta}$ is a consistent estimator of $\theta$ and that $\sqrt{n} (\hat{\theta} - \theta)$ is asymptotically normally distributed with mean zero and variance $\lim_{n \to \infty} \frac{1}{n} \frac{\hat{P}_i^{12}}{\hat{P}_i \hat{Q}_i}$. The existence of this limit is guaranteed by assumption 4.

For compactness, we will rewrite (1) as

$$l_1 = \sum_{i=1}^{n} \hat{r}_{ii} - 0$$

where by definition
Unbiased Estimators

\[ \Gamma_{11} = (u_1 - P_1)P_1'P_1Q_1 \quad \text{(4)} \]

Now \( \hat{L}_1 \) considered as a function of \( \hat{\theta} \) can be expanded formally in powers of \( \hat{\theta} - \theta \), as follows:

\[ \hat{L}_1 = \sum \Gamma_{11} + (\hat{\theta} - \theta) \sum \Gamma_{21} + \frac{1}{2} (\hat{\theta} - \theta)^2 \sum \Gamma_{31} + \ldots \]

where we define

\[ \Gamma_{s1} \equiv \frac{d^s}{d\theta^s} \log P_1'Q_1 \quad (s = 1, 2, \ldots) \quad \text{(5)} \]

This definition is consistent with (3).

Let \( x \equiv \hat{\theta} - \theta \), \( \Gamma_s \equiv \sum \Gamma_{s1} \). Rather than proving the convergence of the power series, let us use a closed form that is always valid:

\[ \hat{L}_1 = \Gamma_1 + x\Gamma_2 + \frac{1}{2} x^2 \Gamma_3 + \frac{1}{6} x^3 \Gamma_4 + \frac{5}{24} x^4 \Gamma_5 \quad \text{(6)} \]

where \( \Gamma_5 \equiv \max \Gamma_5 \) and \( |\delta| < 1 \).

1.2 Derivatives and Expectations

To proceed further, it is necessary to evaluate the \( \Gamma_{k1} \). It is found that
Unbiased Estimators

\[ \Gamma_{ki} = (-A_i)^{k-1} \frac{p_i^k}{Q_i} \sum_{s=1}^{k-1} \mathcal{G}_{k-1}^s \left( \frac{-Q_i}{1 - c_i} \right)^s (-1 + \frac{u_i c_i}{f_i^{s+1}}) \]  

where \( \mathcal{G}_{k-1}^s \) is a Stirling number of the second kind (Jordan, 1947, pp. 31-32, 168).

Define

\[ \gamma_{si} = \delta_{rsi} \]  
\[ \epsilon_{si} = \gamma_{si} - \delta_{rsi} \]  

Since \( \delta u_i = p_i \), we find that

\[ \gamma_{1i} = 0 \]  
\[ \gamma_{2i} = -\frac{p_i^2}{Q_i} \]  
\[ \gamma_{3i} = \frac{A_i^2}{(1 - c_i)^2} \frac{p_i^2}{p_i} (p_i - c_i) \{2(p_i^2 - c_i) - p_i(1 - c_i)\} \]  
\[ \epsilon_{1i} = \gamma_{1i} = (u_i - p_i)p_i^2/Q_i \]  
\[ \epsilon_{2i} = \frac{A_i c_i}{(1 - c_i)} \frac{p_i(u_i - p_i)}{p_i^2} \]  

Let

\[ \gamma_s = \sum_{i} \gamma_{si}/n \quad \epsilon_s = \sum_{i} \epsilon_{si}/n \]
Unbiased Estimators

We will denote the Fisher information by

$$I = -\delta(dL_1/d\theta) = -n\gamma_2 = \sum_i P_i^2 / P_i Q_i$$  \hspace{1cm} (16)

Setting (6) equal to zero, the likelihood equation can now be written in terms of the $\gamma_s$ and the $\epsilon_s$ as

$$-\epsilon_1 = x(\gamma_2 + \epsilon_2) + \frac{1}{2} x^2 (\gamma_3 + \epsilon_3) + \frac{1}{6} x^3 (\gamma_4 + \epsilon_4) + \frac{\delta}{24} x^4 \gamma_5$$  \hspace{1cm} (17)

We will need some information about the order of magnitude of the terms such as those in (17). It may be seen from (7) that each $\epsilon_s$ has the form

$$\epsilon_s = \frac{1}{n} \sum_i K_{si} (u_i - P_i)$$

where $K_{si}$ does not depend on $n$ or on $u_i$. Since $P_i$, $Q_i$ and $1 - c_i$ are bounded, the $K_{si}$ and thus $\epsilon_s$ is bounded. By assumption (4), the bound does not depend on $n$. The same conclusion holds for $\gamma_s$.

Since $\sqrt{n} x$ is asymptotically normally distributed with zero mean and finite variance, it follows that $\delta x^r$ ($r = 1, 2, ...$) is of order $n^{-r/2}$. A similar statement is true of $\sqrt{n} \epsilon_s$. Thus finally $\delta x^r \epsilon^t_s \leq (\delta x^r \delta^t_s)^{1/2}$ so that $\delta x^r \epsilon^t_s$ is of order $n^{-(r+t)/2}$ ($r, t = 1, 2, ...$).
1.3 First-Order Variance of $\hat{\theta}$

To clarify the procedure, let us derive from (17) the familiar formula for the asymptotic variance of $\hat{\theta}$. Square (17) and take expectations to obtain

$$\delta\varepsilon_1^2 = \gamma_2 \delta x^2 + 2 \gamma_2 \delta x^2 \varepsilon_2 + \delta x^2 \varepsilon_2^2 + \gamma_2 \gamma_3 \delta x^2 + \gamma_2 \delta\varepsilon_3^2 + \ldots \quad (18)$$

If we wish to neglect terms $o(n^{-1})$ (of higher order than $n^{-1}$), equation (18) becomes

$$\delta x^2 = \frac{1}{\gamma_2} \delta\varepsilon_1^2 + o(n^{-1}) \quad (19)$$

By (13) and (16), because of local independence,

$$\delta\varepsilon_1^2 = \frac{1}{n^2} \Sigma \frac{P_i}{P_i Q_i} (u_i - P_i) \Sigma \frac{P_j}{P_j Q_j} (u_j - P_j)$$

$$= \frac{1}{n^2} \Sigma \frac{P_i P_j}{P_i Q_i P_j Q_j} \delta(u_i - P_i)(u_j - P_j)$$

$$= \frac{1}{n^2} \Sigma \frac{P_i^2}{P_i Q_i} \text{Var} u_i$$

$$= \frac{1}{n^2} \Sigma \frac{P_i^2}{P_i Q_i}$$

$$= \frac{1}{n^2} \Sigma \frac{P_i^2}{P_i Q_i}$$

$$= \frac{1}{n^2} \quad (20)$$
Unbiased Estimators

Thus, finally

$$\text{Var } \hat{\theta} = \frac{1}{1} + o(n^{-1})$$, \hspace{1cm} (21)

a well-known result. It is derived here to clarify the reasoning to be used subsequently. If \( \hat{\theta} \) is substituted for \( \theta \) on the right side of (21), the formula will still be correct to the specified order of approximation.

1.4 Statistical Bias of \( \hat{\theta} \)

Take the expectation of (17) to obtain

$$-\delta_1 e_1 = \gamma_2 \delta_1 x + \delta_1 xc_2 + \frac{1}{2} \gamma_3 \delta_1 x^2$$, \hspace{1cm} (22)

where \( \delta_1 \) indicates an expectation in which only terms of order \( n^{-1} \) are to be retained. Also multiply (17) by \( c_2 \) and take expectations to obtain

$$-\delta_1 e_1 c_2 = \gamma_2 \delta_1 xc_2$$ \hspace{1cm} (23)

By (9)

$$\delta e_r = 0 \hspace{1cm} r = 1, 2, ...$$ \hspace{1cm} (24)

From (13) and (14)
Unbiased Estimators

\[ \delta \varepsilon_1 \varepsilon_2 = \frac{1}{n} \sum \frac{A_i c_i}{1 - c_i} \frac{P_i^2}{P_{i1}^2} \delta (u_i - P_i)^2 \]

\[ \substack{\text{Substituting (16) and (25) into (23), we have the covariance} \\ \text{covariance}} \]

\[ \delta x_{12} = \frac{1}{n} \sum \frac{A_i c_i}{1 - c_i} \frac{P_i^2}{P_{i1}^2} . \]

Finally, substituting (16), (21), (24), and (26) into (22) and solving for \( \delta x_1 \), we have the bias

\[ B_1(\theta) \equiv \delta 1(\theta - \theta) = \frac{1}{2} \left( \sum \frac{A_i c_i}{1 - c_i} \frac{P_i^2}{P_{i1}^2} + \frac{1}{2} \gamma_3 \right) . \]

This may be rewritten as

\[ B_1(\theta) = \frac{1}{n} \sum_{i=1}^{n} \frac{A_i c_i}{1 - c_i} (A_1 \frac{P_i^2}{P_{i1}^2}) . \]

where

\[ \gamma_1 \equiv \frac{P_i - c_i}{1 - c_i} \text{ and } I_1 \equiv \frac{P_i^2}{P_{i1}^2} . \]
Since $I$ is of order $n$, $B_1(\hat{\theta})$ is of order $n^{-1}$. It may be of interest to note that in the special case where all items are equivalent (all $P_i$ are the same), the bias simplifies to $B_1(\hat{\theta}) = P/nP'$.

1.5 Numerical Results

A hypothetical test was designed to approximate the College Entrance Examination Board's Scholastic Aptitude Test, Verbal Section. This test is composed of $n = 90$ five-choice items. Some information about the distributions of the parameters of the 90 hypothetical items is given in Table 1.

The standard error and bias of $\hat{\theta}$ were computed from (21) and from (27) respectively for various values of $\theta$. The results are shown in Table 2. It appears that the bias in $\hat{\theta}$ is negligible for moderate values of $\theta$, but is sizable for extreme values. Note that the bias is positively correlated with $\theta$. Because of guessing, zero bias does not occur at $\theta = 0$ but at $\theta = .34$ approximately.

1.6 Variance and Bias of Estimated True Score

Since the ability scale is not unique, any monotonic transformation of $\theta$ can serve as a measure of ability. Two transformations are particularly useful: $e^\theta$ and

$$\zeta = \frac{1}{n} \sum_{i=1}^{n} f_i(\theta),$$

(30)

the proportion-correct true score (the number-right true score divided
Unbiased Estimators

TABLE 1

Range and Quartiles of the Item Parameters in 90-Item Hypothetical Test

<table>
<thead>
<tr>
<th></th>
<th>$a_i \cdot A_i / 1.7$</th>
<th>$b_i$</th>
<th>$c_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest value</td>
<td>1.88</td>
<td>2.32</td>
<td>.47</td>
</tr>
<tr>
<td>$Q_1$</td>
<td>1.07</td>
<td>1.15</td>
<td>.20</td>
</tr>
<tr>
<td>Median</td>
<td>.83</td>
<td>.38</td>
<td>.15</td>
</tr>
<tr>
<td>$Q_3$</td>
<td>.69</td>
<td>-.41</td>
<td>.13</td>
</tr>
<tr>
<td>Lowest value</td>
<td>.41</td>
<td>-3.94</td>
<td>.01</td>
</tr>
</tbody>
</table>
by the number of items). One important reason for using the latter transformation is the following.

Ordinarily, as in Table 2, we find large standard errors of \( \theta \) where \( \theta \) is extreme. Usually these large standard errors are no more harmful to the user than are the smaller standard errors found when \( \theta \) is near the level aimed at by the test. There is a reason why this is so: If it were not, the user should have designed his test so as to reduce those standard errors that were troublesome to him.

We see that from this point of view the size of a difference on the \( \theta \) scale does not correspond to its importance. The discrepancy is greatly reduced, however, if we measure ability on the \( \zeta \) scale instead of on the \( \theta \) scale. This is one reason, among several, why we are interested in the variance and bias of

\[
\zeta = \sum_{i=1}^{n} p_i(\hat{\theta})/n
\]  

(31)

Although the proportion-correct true score

\[
z = \sum_{i=1}^{n} u_i/n
\]  

(32)

is an unbiased estimator of \( \zeta \), \( z \) is never a fully efficient estimator of \( \zeta \) unless \( c_i = 0 \) and \( a_i = a_j \) \((i,j = 1,2,\ldots,n)\):

the sampling variance

\[
\text{Var } z = \frac{1}{n^2} \sum_{i=1}^{n} p_i q_i
\]  

(33)
<table>
<thead>
<tr>
<th>$\hat{\theta}$</th>
<th>$\sqrt{\text{Var}\hat{\theta}}$</th>
<th>$B(\hat{\theta})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>.60</td>
<td>.24</td>
</tr>
<tr>
<td>3.0</td>
<td>.43</td>
<td>.12</td>
</tr>
<tr>
<td>2.5</td>
<td>.31</td>
<td>.06</td>
</tr>
<tr>
<td>2.0</td>
<td>.23</td>
<td>.032</td>
</tr>
<tr>
<td>1.5</td>
<td>.19</td>
<td>.011</td>
</tr>
<tr>
<td>1.0</td>
<td>.19</td>
<td>.0032</td>
</tr>
<tr>
<td>0.5</td>
<td>.20</td>
<td>.0012</td>
</tr>
<tr>
<td>0</td>
<td>.22</td>
<td>-.0028</td>
</tr>
<tr>
<td>-.5</td>
<td>.25</td>
<td>-.010</td>
</tr>
<tr>
<td>-1.0</td>
<td>.31</td>
<td>-.025</td>
</tr>
<tr>
<td>-1.5</td>
<td>.41</td>
<td>-.05</td>
</tr>
<tr>
<td>-2.0</td>
<td>.54</td>
<td>-.09</td>
</tr>
<tr>
<td>-2.5</td>
<td>.70</td>
<td>-.14</td>
</tr>
<tr>
<td>-3.0</td>
<td>.89</td>
<td>-.22</td>
</tr>
<tr>
<td>-3.5</td>
<td>1.09</td>
<td>-.31</td>
</tr>
</tbody>
</table>
is not as small as the sampling variance of \( \hat{\zeta} \), which we must now derive.

By (31)

\[
d\hat{\zeta} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j} \, dP_{i,j} \delta_{ij} .
\]  

(34)

Using the 'delta' method

\[
\text{Var} \hat{\zeta} = \frac{1}{n^2} \left( \sum_{i=1}^{n} P_{i} \right)^2 \text{Var} \hat{\theta} .
\]

By (21) and (16)

\[
\text{Var} \hat{\zeta} = \frac{1}{n^2} \frac{\left( \sum_{i=1}^{n} P_{i} \right)^2}{\frac{1}{P_{i}^2} \sum_{i=1}^{n} \frac{1}{P_{i} Q_{i}}} .
\]

(35)

To find the bias of \( \hat{\zeta} \), we expand it in powers of \( x = \zeta - \hat{\zeta} \):

\[
\hat{\zeta} - \zeta = \frac{x}{n} \sum_{i=1}^{n} P_{i} + \frac{x^2}{2n} \sum_{i=1}^{n} P_{i}^2 + \ldots
\]

(36)
where

\[ p''_1 = \frac{d^2 p_1}{d \phi^2} \]

Taking expectations, and neglecting higher-order terms, we have for the bias

\[ B_1(\hat{\phi}) = \delta(\hat{\phi} - \phi) = \frac{1}{n} [B(\phi) \Sigma p'_1 + \frac{1}{2} \text{Var } \hat{\phi}(p''_1)] \]  \hspace{1cm} (37)

This can be rewritten as

\[ B_1(\hat{\phi}) = \frac{\zeta'}{2} \left( \sum \frac{A_i c_i p''_1^2}{(1 - c_i p''_1)} + \frac{1}{2} \sum \gamma_{31} \right) + \frac{\zeta''}{2I} \]  \hspace{1cm} (38)

where \( \zeta' = \Sigma p'_1/n \) and \( \zeta'' = \Sigma p''_1/n \). Let us note is passing that when all items are equivalent (all \( p'_1(\phi) \) are the same), \( \hat{\phi} = z \) and its bias (38) is zero.

1.7 Numerical Results

Table 3 shows the bias in \( \zeta \) for the same hypothetical test considered in Section 1.5. The biases are all positive. However, they are negligible at all except the lowest ability levels. This tends to confirm our choice of the \( \zeta \) scale of ability rather than the \( \theta \) scale for many purposes.
### Unbiased Estimators

#### TABLE 3

Standard Error of $z$ and of $\hat{\zeta}$, and Statistical Bias of $\hat{\zeta}$

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\hat{\zeta}$</th>
<th>$\sqrt{\text{Var} \ z}$</th>
<th>$\sqrt{\text{Var} \ \hat{\zeta}}$</th>
<th>$B(\hat{\zeta})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>0.981</td>
<td>0.014</td>
<td>0.014</td>
<td>0.00045</td>
</tr>
<tr>
<td>3.0</td>
<td>0.966</td>
<td>0.019</td>
<td>0.018</td>
<td>0.00052</td>
</tr>
<tr>
<td>2.5</td>
<td>0.937</td>
<td>0.024</td>
<td>0.023</td>
<td>0.00064</td>
</tr>
<tr>
<td>2.0</td>
<td>0.891</td>
<td>0.031</td>
<td>0.029</td>
<td>0.00059</td>
</tr>
<tr>
<td>1.5</td>
<td>0.812</td>
<td>0.037</td>
<td>0.035</td>
<td>0.00021</td>
</tr>
<tr>
<td>1.0</td>
<td>0.715</td>
<td>0.042</td>
<td>0.040</td>
<td>0.00026</td>
</tr>
<tr>
<td>0.5</td>
<td>0.608</td>
<td>0.045</td>
<td>0.042</td>
<td>0.00061</td>
</tr>
<tr>
<td>0</td>
<td>0.506</td>
<td>0.046</td>
<td>0.043</td>
<td>0.00061</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.416</td>
<td>0.047</td>
<td>0.042</td>
<td>0.00062</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.344</td>
<td>0.046</td>
<td>0.038</td>
<td>0.00061</td>
</tr>
<tr>
<td>-1.5</td>
<td>0.291</td>
<td>0.045</td>
<td>0.037</td>
<td>0.00085</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.254</td>
<td>0.044</td>
<td>0.033</td>
<td>0.0014</td>
</tr>
<tr>
<td>-2.5</td>
<td>0.227</td>
<td>0.042</td>
<td>0.029</td>
<td>0.0020</td>
</tr>
<tr>
<td>-3.0</td>
<td>0.211</td>
<td>0.042</td>
<td>0.025</td>
<td>0.0024</td>
</tr>
<tr>
<td>-3.5</td>
<td>0.199</td>
<td>0.041</td>
<td>0.021</td>
<td>0.0026</td>
</tr>
</tbody>
</table>
As a matter of incidental interest, for selected values of true score, Table 3 compares the standard error (35) of the maximum-likelihood estimator \( \hat{\zeta} \) with the standard error (33) of the unbiased estimator \( z \) (proportion-correct score). There is little difference in accuracy between the two estimators for \( \zeta > .5 \). At low true-score levels, the maximum-likelihood estimator is much better than the proportion of correct answers.

2. Unbiased Estimation of \( s_\theta^2 \), of \( s_\zeta^2 \); Test Reliability

The symbols \( s_\theta^2 \) and \( s_\zeta^2 \) are used for the sample variance of \( \theta \) and of \( \zeta \) across the \( N \) examinees in the sample:

\[
s_\theta^2 = \frac{1}{N} \sum_{a=1}^{N} \theta_a^2 - \left( \frac{1}{N} \sum_{a=1}^{N} \theta_a \right)^2 .
\]

The maximum-likelihood estimators of \( s_\theta^2 \) and \( s_\zeta^2 \) are \( s_\theta^2 \) and \( s_\zeta^2 \), the sample variances across examinees of \( \theta \) and of \( \zeta \).

2.1 Asymptotically Unbiased \( \hat{\theta} \) Estimator of \( \sigma_\theta^2 \)

Assume that our examinees are a random sample of \( N \) from some population. Denote by \( \sigma_\theta^2 \) the population variance of \( \theta \). Then \( Ns_\theta^2/(N-1) \) is an unbiased estimator of \( \sigma_\theta^2 \). Since \( s_\theta^2 \) is unobservable, our first task is to find a function of \( \theta \) that is an asymptotically unbiased estimator of \( \sigma_\theta^2 \).
By the formula for the variance of a sum we have

\[ \sigma^2_{\theta} \equiv \sigma^2_\theta + \sigma^2_x + 2\sigma_{\theta x}, \quad (40) \]

where \( \sigma^2 \) denotes a variance across all examinees in the population and \( \sigma_{\theta x} \) is the corresponding population covariance. By a well-known identity from the analysis of variance

\[ \sigma^2_x \equiv \mathbb{E}^2(x|\theta) + \sigma^2_\theta \cdot (x|\theta), \quad (41) \]

where \( \mathbb{E}_\theta \) denotes an expectation across all examinees in the population. Similarly,

\[ \sigma_{\theta x} \equiv \mathbb{E}_\theta (x|\theta). \quad (42) \]

Substituting (41) and (42) into (40), transposing, writing \( B_1 \equiv \mathbb{E}_1(x|\theta) \) as in (28), and dropping the subscript from \( B_1 \) for convenience, we have

\[ \sigma^2_{\theta} \equiv \sigma^2_\theta - 2\sigma_{\theta B} - \mathbb{E}^2_\theta(x|\theta) - \sigma^2_B. \quad (43) \]

Since by (28) \( B \) is of order \( n^{-1} \), its variance is of order \( n^{-2} \), so \( \sigma^2_B \) can be neglected in (43). Since Section 1 deals with a single fixed examinee, the symbol \( \sigma^2_{x|\theta} \) in (43) has the same meaning as \( \text{Var} \theta \) in (21):

\[ \sigma^2_{x|\theta} = \frac{1}{I(\theta)} + o(n^{-1}) \]
Unbiased Estimators

21

where $I = I(\theta)$ is given by (16). Since $\sigma_x^2|\theta$ is of order $n^{-1}$, the
effect of replacing $\theta$ by $\hat{\theta}$ on the right is negligible:

$$\delta_0 \sigma_x^2|\theta = \delta_0 \frac{1}{I(\theta)} + o(n^{-1}) .$$

By similar reasoning, we may replace $\sigma_{\theta B}$ in (43) by $\sigma_{\theta B}^\wedge$ where $\hat{B}$
is defined by (27) with $\theta$ replaced by $\hat{\theta}$. The result of these
approximations is that

$$\sigma_\theta^2 = \sigma_\theta^2 - 2\sigma_{\theta B}^\wedge - \delta_0 \frac{1}{I(\hat{\theta})} + o(n^{-1}) . \quad (44)$$

A useful estimator of $\sigma_\theta^2$ can be calculated from

$$\sigma_\theta^2 = \frac{N}{N-1} \hat{s}_\theta^2 - \frac{2N}{N-1} \hat{s}_{\theta B}^\wedge - \frac{1}{N} \frac{1}{a=1} \frac{1}{I(\hat{\theta}_a)} , \quad (45)$$

where

$$\hat{s}_{\theta B}^\wedge = \frac{1}{N} \frac{1}{a=1} \hat{\theta}_a \hat{B}_a - \left( \frac{1}{N} \frac{1}{a=1} \hat{\theta}_a \right) \left( \frac{1}{N} \frac{1}{a=1} \hat{B}_a \right)$$

and $\hat{B}_a$ is given by (27) with $\theta$ replaced by $\hat{\theta}_a$. If we wish to
estimate the sample variance of ability $s_\theta^2$ rather than the population
variance $\sigma_\theta^2$, we can use

$$s_\theta^2 = s_\theta^2 - 2\hat{s}_{\theta B}^\wedge - \frac{N-1}{N^2} \frac{1}{a=1} \frac{1}{I(\hat{\theta}_a)} . \quad (46)$$
The second and third terms of (44) are of order \( n^{-1} \), an order of magnitude smaller than the first term but larger than the neglected terms. The covariance of \( \hat{\theta} \) and \( \hat{B} \) is usually positive, as can be readily seen from Table 2. Since \( I(\hat{\theta}) \) is necessarily positive, it appears that usually \( \sigma_\theta^2 < \sigma_\theta^2 \), an inequality that is frequently assumed without proof. It is not clear whether this inequality is necessarily true.

2.2 The Reliability of \( \hat{\theta} \)

Consider the parallel-forms reliability coefficient \( \rho_{\theta\theta'} \), the correlation between scores \( \hat{\theta} \) and \( \hat{\theta}' \) on two parallel tests. For present purposes, two tests are parallel when for each item in one test there is an item in the other test with the same item response function. Let us estimate

\[
\rho_{\theta\theta'} = \frac{\sigma_{\theta\theta}^2}{\sigma_\theta^2} = \frac{\sigma_{\theta\theta}^2}{\sigma_\theta^2}
\]

(47)

from a single test administration by substituting asymptotically unbiased estimators of the numerator and of the denominator into (47).

As in (41),

\[
\sigma_{\theta\theta'} = E[\theta \theta' | \theta] + E[\theta |\theta], E(\delta | \theta)
\]

(48)
Because of local independence, the first term on the right vanishes.

Because of parallelism, the two expectations in the last term are identical, so this term is a variance. We thus have

\[ \sigma_{\theta\theta}^2 = \sigma_\delta^2(\delta|\theta) = \sigma_B^2 + \sigma_\theta^2 + 2\sigma_{B\theta} \tag{49} \]

From (49) and (43),

\[ \sigma_{\theta\theta}^2 = \sigma_\theta^2 - \delta_\theta^2 \sigma_x|\theta \tag{50} \]

We see that the parallel-forms reliability of \( \hat{\theta} \) is

\[ \rho_{\theta\theta} = 1 - \frac{1}{\sigma_\theta^2} \frac{1}{I(\theta)} + o(n^{-1}) \tag{51} \]

Priority in obtaining this result belongs to Symson [Note 1].

Replacing population values on the right by the corresponding sample statistics, we have a sample estimator of the parallel-forms reliability coefficient of \( \hat{\theta} \):

\[ \hat{\rho}_{\theta\theta} = \frac{N - 1}{N} \sum_{a=1}^{N} \frac{1}{I(\theta_a)} \tag{52} \]

Since \( \hat{\theta} \) is neither unbiased nor uncorrelated with \( \theta \), we should not expect the usual reliability formulas of classical test theory to apply. A similar but not identical case is discussed in
Unbiased Estimators

Lord and Novick (1968, Section 9.8). Thus \( \rho_{\hat{\theta} \hat{\theta}} \), \( \rho_{\theta \theta}^2 \), and \( \sigma_{\hat{\theta}}^2 / \sigma_{\theta}^2 \) are not interchangeable definitions of reliability. Since correlational measures are hard to interpret in the absence of linearity and homoscedasticity, we will not now push this investigation of reliability further.

2.3 Corresponding Results for True Score

By the same reasoning used to obtain (44) we have

\[
\sigma^2_{\zeta} = \sigma^2_{\zeta} - 2 \sigma_{\zeta, B(\zeta)} - \delta_{\theta} \sigma^2_{\zeta} \| \zeta - \sigma^2_{B(\zeta)}
\]

\[
= \sigma^2_{\zeta} - 2 \sigma_{\zeta, B(\zeta)} - \delta_{\theta} \frac{\zeta^2}{I(\hat{\theta})} + o(n^{-1}) .
\]\n
(53)

A useful estimator of \( \sigma^2_{\zeta} \) can be calculated from

\[
\sigma^2_{\zeta} \equiv \frac{N}{N - 1} \frac{2}{N - 1} \sigma^2_{\zeta, B(\zeta)} - \frac{1}{N} \frac{N}{\Sigma_{a=1}^{\infty}} \frac{\zeta^2_a}{I(\theta_a)}
\]\n
(54)

To estimate \( \sigma^2_{\zeta} \), we can use

\[
\hat{s}^2_{\zeta} \equiv \hat{s}^2_{\zeta} - 2 \sigma_{\zeta, B(\zeta)} - \frac{N - 1}{N} \sum_{a=1}^{\infty} \frac{\zeta^2_a}{I(\theta_a)}
\]\n
(55)
As in (50) - (52) we have

\[ \sigma_{\theta}^2 = \sigma_{\theta}^2 - \delta_{\theta} \sigma_{\theta}^2 / \zeta \]  

(56)

\[ \rho_{\theta}^2 = 1 - \frac{1}{2} \delta_{\theta} \frac{\hat{\theta}^2}{I(\theta)} + o(n^{-1}) \]  

(57)

\[ \rho_{\theta}^2 = 1 - \frac{N-1}{2 \sigma_{\theta}^2} \sum_{a=1}^{N} \frac{\hat{\theta}_{a}^2}{I(\hat{\theta}_{a})} \]  

(58)

2.4 Numerical Results for True Scores

At moderate ability levels, (28) provides adequate but usually negligible corrections for bias in \( \hat{\theta} \). Experience shows that at very low ability levels, the usual test length (\( n \)) of 50 or 100 items is not long enough for the asymptotic results of (28) to apply. For example, an examinee whose true \( \theta \) is -3 may easily obtain an estimated ability \( \hat{\theta} \) of -30 or of \( -\infty \). For sufficiently long tests, such extreme values of \( \hat{\theta} \) would have negligible probability, but with the usual values of \( n \) and equation (28) is totally inadequate for correcting \( \hat{\theta} \) for bias at low ability levels.

This same difficulty carries over to the unbiased estimation of \( \sigma^2_{\theta} \) using (46). Since all ability levels are involved in (46), the formula is useless in practice for any group that contains even a few low-ability examinees. Fortunately, this difficulty does not
carry over to the estimation of ability on the true-score ( ) scale.

The hypothetical SAT Verbal Test of Tables 1-3 was administered to a typical group of 2995 hypothetical examinees. The bias in was estimated for each examinee and a corrected obtained from (51):

\[
\text{corrected } \hat{\xi} = \hat{\xi} - B_1(\hat{\xi})
\]

In a few cases where the corrected would have been below the chance level \( \sum_1^n c_i \), the corrected was set equal to \( \sum_1^n c_i \).

The mean of the 2995 true \( \xi \) used to generate the data was .5280, the mean of the uncorrected \( \hat{\xi} \) was .5294, the mean of the corrected \( \hat{\xi} \) was .5288. Thus the correction was in the right direction, but not large enough. The uncorrected mean \( \hat{\xi} \) was already so accurate as to leave little room for improvement.

Next, (55) was used to estimate \( s_\xi \). The true value was \( s_\xi = .1610 \), the standard deviation of \( \hat{\xi} \) was \( s_\xi = .1660 \), the corrected estimate from (55) was \( \hat{s}_\xi = .1614 \). The correction worked very well here.

The parallel-forms reliability of \( \hat{\xi} \) was estimated from (58) to be \( \rho_{\xi\xi} = .9420 \). We have no 'true' value against which this can be compared, but the estimate seems a reasonable one. The Kuder-Richardson formula-20 reliability of number-right scores for these data is .9275.
It should be remembered that both the formulas and the numerical results in this report apply in situations where the item parameters are known. These formulas may be satisfactory for situations where the item parameters have been estimated from large groups not containing the examinees whose ability estimates are to be corrected for bias. These formulas will not be adequate for situations where the item parameters and ability parameters are estimated simultaneously from a single data set.
Reference Note

References


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

6 Commanding Officer  
Naval Research Laboratory  
Code 2627  
Washington, DC 20390

1 Psychologist  
ONR Branch Office  
Building 114, Section D  
666 Summer Street  
Boston, MA 02210

1 Office of Naval Research  
Code 437  
800 North Quincy Street  
Arlington, VA 22217

5 Personnel and Training Research Programs  
Code 458  
Office of Naval Research  
Arlington, VA 22217

1 Psychologist  
ONR Branch Office  
1030 East Green Street  
Pasadena, CA 91101

1 Office of the Chief of Naval Operations  
Research Development and Studies Branch  
OP-113  
Washington, DC 20350

1 LT Frank C. Pecho, MSC, USN (Ph.D.)  
Selection and Training Research Division  
Human Performance Sciences Department  
Naval Aerospace Medical Research Lab.  
Pensacola, FL 32508

1 Dr. Bernard Rimland (038)  
Navy Personnel R & D Center  
San Diego, CA 92152

1 Mr. Arnold Rubenstein  
Office of Naval Technology  
800 N. Quincy Street  
Arlington, VA 22217

1 Dr. Worth Scanland, Director  
Research, Development, Test and Evaluation  
N-5  
Naval Education and Training Command  
NAS  
Pensacola, FL 32508

1 Dr. Robert G. Smith  
Office of Chief of Naval Operations  
OP-987H  
Washington, DC 20350

1 Dr. Alfred F. Spade  
Training Analysis and Evaluation Group  
Department of the Navy  
Orlando, Fl. 32813

1 Dr. Richard Sorensen  
Navy Personnel R & D Center  
San Diego, CA 92152

1 Mr. J. B. Sympton  
Navy Personnel R & D Center  
San Diego, CA 92152

1 Dr. Ronald Weitman  
Code 34 WZ  
Department of Administrative Services  
U.S. Naval Postgraduate School  
Monterey, CA 93940

1 Dr. Robert Wherry  
562 Mallard Drive  
Chalfont, PA 18914
DISTRIBUTION LIST

Navy

1 Dr. Ed Aiken
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Jack R. Borsting
Provost and Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code H-711
NAVTRADEQUIPCEN
Orlando, FL 32813

1 Chief of Naval Education and Training Liaison Office
Air Force Human Resource Laboratory
Flying Training Division
Williams Air Force Base, AZ 85214

1 CDR Mike Curran
Office of Naval Research
600 North Quincy Street
Code 270
Arlington, VA 22217

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

1 Dr. Pat Federico
Navy Personnel R & D Center
San Diego, CA 92152

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

1 Dr. John Ford
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Patrick R. Harrison
Psychology Course Director
Leadership and Law Department (7b)
Division of Professional Development
U.S. Naval Academy
Annapolis, MD 21402

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. William L. Maley
Principal Civilian Advisor for Education and Training
Naval Training Command, Code OOA
Pensacola, FL 32508

1 Dr. Kneale Marshall
Scientific Advisor to DCNO(MPT)
OpUIt
Washington, DC 20370

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

1 Dr. William Montague
Navy Personnel R & D Center
San Diego, CA 92152

1 Mr. William Nordbrock
Instructional Program Development
Building 90
MST-PDCC
Great Lakes NTC, IL 60088

1 Library, Code P201L
Navy Personnel R & D Center
San Diego, CA 92152
1 Dr. Robert Wisher  
Code 309  
Navy Personnel R & D Center  
San Diego, CA 92152

1 Dr. Martin F. Wiskoff  
Navy Personnel R & D Center  
San Diego, CA 92152

1 Mr. Ted M. I. Yellen  
Technical Information Office  
Code 201  
Navy Personnel R & D Center  
San Diego, CA 92152

Army

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

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

1 Dr. Dexter Fletcher  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 COL Frank Hart  
Army Research Institute for the  
Behavioral & Social Sciences  
5001 Eisenhower Blvd.  
Alexandria, VA 22333

1 Dr. Michael Kaplan  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Dr. Robert Sasmor  
San Diego, CA 92152  
U.S. Army Research Institute for the Social and Behavioral Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Dr. Milton S. Katz  
Training Technical Area  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.  
Attn: PERI-OK  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

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

1 Commandant  
U.S. Army Institute of Administration  
Attn: Dr. Sherrill  
Ft. Benning NN, IN 46256

1 Dr. Frederick Steinheiser  
Department of the Navy  
Chief of Naval Operations  
OP-113  
Washington, DC 20350

1 Dr. Joseph Ward  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Laboratory  
AFRHL/HFD  
Brooks Air Force Base, TX 78235

1 U.S. Air Force Office of  
Scientific Research  
Life Sciences Directorate  
Bolling Air Force Base  
Washington, DC 20332
| Military Assistant for Training and Personnel Technology | Dr. Joseph L. Young, Director |
| Office of the Under Secretary of Defense for Research and Engineering | Memory and Cognitive Processes |
| Room 3D129, The Pentagon | National Science Foundation |
| Washington, DC 20550 | Washington, DC 20550 |

| Dr. Wayne Sellman | Civil Government |
| Office of the Assistant Secretary of Defense (MARAD) | |
| 28209 The Pentagon | |
| Washington, DC 20301 | |

| Dr. Susan Chipman | 1 |
| Learning and Development | Dr. Alan Baddeley |
| National Institute of Education | Medical Research Council |
| 1400 19th Street, NW | Applied Psychology Unit |
| Washington, DC 20030 | 15 Chaucer Road |
| | Cambridge CB2 2EF |
| | ENGLAND |

| Dr. Arthur Melmed | 1 |
| National Institute of Education | Dr. Isaac Bejer |
| 1200 19th Street, N.W. | Educational Testing Service |
| Washington, DC 20036 | Princeton, NJ 08541 |

| Dr. Andre R. Holm | 1 |
| Science Education Development and Research | Dr. Manoucheh Birenbaum |
| National Science Foundation | School of Education |
| Washington, DC 20550 | Tel Aviv University |
| | Tel Aviv, Ramat Aviv 69976 |
| | ISRAEL |

| Dr. Vern W. Urry | 1 |
| Personnel R & D Center | Dr. Werner Birke |
| Office of Personnel Management | WesdPE im Streitkräfteamt |
| 1900 E Street, NW | Postfach 20 503 |
| Washington, DC 20415 | U-5300 Bonn 2 |
| | WEST GERMANY |
Dr. Michael Levine
Department of Educational Psychology
210 Education Building
University of Illinois
Champaign, IL 61801

Dr. Melvin A. Novick
J56 Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

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

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army Navy Drive
Arlington, VA 22202

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

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

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

Dr. James R. Panferlo
2431 North Edgewood Street
Arlington, VA 22207

Dr. Diane H. Ramsey-Klee
K-12 Research and System Design
3947 Ridgemont Drive
Malibu, CA 90265

Mr. Minrat H. L. Rauch
P 11 4
Bundesministerium der Verteidigung
Postfach 1326
D-53 Bonn 1
GERMANY

Dr. Mark D. Reckase
Educational Psychology Department
University of Missouri-Columbia
4 Hill Mall
Columbia, MO 65211

Dr. Gary Marco
Educational Testing Service
Princeton, NJ 08541

Dr. Scott Maxwell
Department of Psychology
University of Houston
Houston, TX 77004

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

Dr. Allen Munro
Behavioral Technology Laboratories
1645 Elena Avenue
Fourth Floor
Redondo Beach, CA 90277
Dr. Robert Tsutakawa  
Department of Statistics  
University of Missouri  
Columbia, MO 65201

Dr. David Vale  
Assessment Systems Corporation  
2395 University Avenue  
Suite 306  
St. Paul, MN 55114

Dr. Howard Wainer  
Educational Testing Service  
Princeton, NJ 08541

Dr. Thomas Wallsten  
Psychometric Laboratory  
Davis Hall 013A  
University of North Carolina  
Chapel Hill, NC 27514

Dr. Phyllis Weaver  
Graduate School of Education  
Harvard University  
200 Larsen Hall, Appian Way  
Cambridge, MA 02138

Dr. David J. Weiss  
Hool Elliott Hall  
University of Minnesota  
75 East River Road  
Minneapolis, MN 55455

Dr. Susan E. Whitely  
Psychology Department  
University of Kansas  
Lawrence, KS 66044

Dr. Wolfgang Wildgrube  
Streirkreisamt  
Box 2U 30 03  
D-5360 Bonn 2  
WEST GERMANY