An Investigation of Methods for Reducing Sampling Error in Certain IRT Procedures

Wingersky, Marilyn S.; Lord, Frederic M.
Educational Testing Service, Princeton, N.J.

Aug 83
N00014-80-C-0402
57p.

Reports - Research/Technical (143)

*Error of Measurement; Estimation (Mathematics); Item Banks; *Latent Trait Theory; Mathematical Formulas; *Maximum Likelihood Statistics; *Research Methodology; Sample Size; *Sampling; Statistical Distributions; Test Length; Validity

The sampling errors of maximum likelihood estimates of item-response theory parameters are studied in the case where both people and item parameters are estimated simultaneously. A check on the validity of the standard error formulas is carried out. The effect of varying sample size, test length, and the shape of the ability distribution is investigated. Finally, the effect of anchor-test length on the standard error of item parameters is studied numerically for the situation, common in equating studies, where two groups of examinees each take a different test form together with the same anchor test. The results encourage the use of rectangular or bimodal ability distributions, also the use of very short anchor tests. (Author)
AN INVESTIGATION OF METHODS FOR REDUCING SAMPLING ERROR IN CERTAIN IRT PROCEDURES

Marilyn S. Wingersky
and
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

August 1983

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

Approved for public release; distribution unlimited.
An Investigation of Methods for Reducing Sampling Error in Certain IRT Procedures

Marilyn S. Wingersky and Frederic M. Lord

Educational Testing Service - Princeton, New Jersey 08541

Office of Naval Research (Code 458)

Arlington, VA 22217

Approved for public release; distribution unlimited.

The sampling errors of maximum likelihood estimates of item-response theory parameters are studied in the case where both people and item parameters are estimated simultaneously. A check on the validity of the standard error formulas is carried out. The effect of varying sample size, test length, and the shape of the ability distribution is investigated. Finally, the effect of anchor-test length on the standard error of item parameters is studied numerically for the situation, common in equating studies, where two groups of...
examinees each take a different test form together with the same anchor test. The results encourage the use of rectangular or bimodal ability distributions, also the use of very short anchor tests.
An Investigation of Methods for Reducing Sampling Error in Certain IRT Procedures

Abstract

The sampling errors of maximum likelihood estimates of item-response theory parameters are studied in the case where both people and item parameters are estimated simultaneously. A check on the validity of the standard error formulas is carried out: The effect of varying sample size, test length, and the shape of the ability distribution is investigated. Finally, the effect of anchor-test length on the standard error of item parameters is studied numerically for the situation, common in equating studies, where two groups of examinees each take a different test form together with the same anchor test. The results encourage the use of rectangular or bimodal ability distributions, also the use of very short anchor tests.
An investigation of Methods for Reducing Sampling Error in Certain IRT Procedures

In IRT until now, the sampling variances and covariances for maximum likelihood estimates of item parameters have usually been computed by assuming the abilities to be known; the sampling variances and covariances for ability estimates were computed by assuming the item parameters to be known. In this paper, a suggested method for computing the sampling variance-covariance matrix when all parameters are unknown (Lord and Wingersky, 1983) will be used to try to answer various practical questions. Section 2 presents needed additional, though not conclusive, evidence that the new method for computing the variance-covariance matrix yields correct results. Section 3 investigates the effect of changing the number of items or the number or distribution of people on the standard errors of the item parameters and of the abilities. Section 4 presents a technique for displaying and understanding the standard errors and sampling covariances of estimates of item parameters.

Section 5 deals with the practically important situation where we have two tests that contain a set of items in common and these tests are administered to two separate groups of examinees. A problem in item

*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.
banking or test equating is putting the parameter estimates for the two tests on a common scale. One way to do this is to estimate all of the parameters for both tests in one calibration run. When this is done, how does the number and quality of the common items affect the standard errors of the parameter estimates for the unique (noncommon) items?

1. Preliminaries

The three-parameter Birnbaum logistic model is used throughout. The probability of examinee $a$ answering item $i$ correctly is

$$P_{ia} = c_i + (1 - c_i)/(1 + \exp(-1.7a_i(\Theta_a - b_i)))$$

where $a_i$ is the discrimination of item $i$, $b_i$ is the difficulty for the item, $c_i$ is the lower asymptote of the item response function, and $\Theta_a$ is the ability for examinee $a$. In a typical calibration run, poorly estimatable $c_i$ are ordinarily fixed at some common value. In this paper, however, all $c_i$ are considered unknown and must be estimated. In treating all of the $c_i$ as unknown we are looking at the "worst case" standard errors.

In IRT, the origin and unit of measurement of the ability scale is arbitrary. Until this scale is specified all parameters except the $c_i$ are unidentifiable. The origin and unit of the ability scale must be specified in terms of (as a function of) the true parameters. If the
origin and unit of the ability scale were specified in terms of the parameter estimates, then the true parameters would be undefined. Since the true parameters are unknown but depend on the scale used, this means that the scale origin and the scale unit (each defined as a function of the true parameters) must be estimated from the data. The estimated origin and scale unit are obviously subject to sampling errors, which affect the accuracy of all parameter estimates. It is therefore important to define the origin and unit each by a function of parameters that can be estimated with good accuracy.

The scale recommended in Lord and Wingersky (1983) and used here requires that the mean of the difficulty parameters of certain selected items be 0 (the origin) and that the difference between two such means for two sets of selected items be 1 (the scale unit). This scale will be referred to as the "capital" scale; parameters on this scale will be denoted by the capital letters $A_i$, $B_i$, $C_i$, $E_a$. The "small" scale or the "LOGIST" scale, referred to by lower-case letters, is the scale used by the LOGIST program (Wingersky, Barton, and Lord (1982)), the computer program used here for estimating the parameters of (1) by maximum likelihood. LOGIST sets a truncated mean of the estimated abilities to 0 and a truncated standard deviation of the estimated abilities to 1. The following formulas convert the parameters from the LOGIST scale to the capital scale:

$$\theta_a = (\theta_a - \bar{\theta}_0) / k$$

$$k = B_1 - B_0$$
where $b_0$ and $b_1$ are means of the $b_i$ for two selected subsets of items. The capital scale is a linear transformation of the LOGIST scale. The $c_i$ are not affected by the scale.

2. Variance of $p_1$, the Proportion Correct

If we could prove that the maximum likelihood parameter estimates for the Birnbaum model are consistent when all item and ability parameters are estimated simultaneously, the sampling variance-covariance matrix described in Lord and Wingersky (1983) would be the correct one to use. Since consistency has not yet been proven mathematically any results that confirm the appropriateness of this variance-covariance matrix makes one feel more comfortable about using it.

The sampling variance of $p_1$, the proportion of examinees in the sample who answer item $i$ correctly, can be computed directly from familiar standard formulas; it can also be computed with some effort from the sampling variance-covariance matrix obtained by Lord and Wingersky (1983). These two methods should give the same results if the Lord-Wingersky matrix is correct.
The usual likelihood equations for $\hat{b}_i$ and for $\hat{c}_i$, obtained by setting the derivative of the likelihood function equal to zero, are (Lord, 1980, eq. 12.1 and 12.2)

\[ \sum_{a=1}^{N} (u_{ia} - \hat{P}_i(\theta_a))(\hat{P}_i(\theta_a) - \hat{c}_i)\hat{P}_i(\theta_a) = 0, \]  

\[ \sum_{a=1}^{N} (\hat{u}_{ia} - \hat{P}_i(\theta_a))\hat{P}_i(\theta_a) = 0, \]  

where $u_{ia}$ is the score (0 or 1) of examinee $a$ on item $i$, $N$ is the number of examinees, and a caret denotes substitution of parameter estimates for true parameter values. Multiplying (3) by $c_i$, adding to (2), and transposing gives

\[ \sum_{a=1}^{N} \hat{P}_i(\theta_a) = \sum_{a=1}^{N} u_{ia}. \]  

Since

\[ P_i = \frac{1}{N} \sum_{a=1}^{N} u_{ia}, \]  

we have

10
\[ p_i = \frac{1}{N} \sum_{a=1}^{N} p_i(\theta_a) \]  

(5)

From (4) and (5), we can derive two separate formulas for the variance of \( p_i \).

For some group of examinees whose abilities are specified by the vector \( \theta = (\theta_1, \theta_2, \ldots, \theta_N) \), we have from (4) that

\[ \text{var}(p_i | \theta) = \frac{1}{N^2} \sum_{a=1}^{N} \sum_{a'=1}^{N} \text{cov}(u_{ia}, u_{ia'}) | \theta \]  

(6)

\[ = \frac{1}{N^2} \sum_{a=1}^{N} \text{var}(u_{ia} | \theta) \]  

\[ = \frac{1}{N^2} \sum_{a=1}^{N} p_i(\theta_a) Q_i(\theta_a) \]  

(6)

with

\[ Q_i(\theta_a) = 1 - p_i(\theta_a) \]

since \( \text{cov}(u_{ia}, u_{ia'} | \theta) = 0 \) when \( a \neq a' \). Similarly,

\[ \text{cov}(p_i, p_j | \theta) = 0 \]  

(7)
By the formula for the covariance between two sums, we have from (5) for the same group of examinees that

\[
\text{var}(p_i | \theta) = \frac{1}{N^2} \sum_{a=1}^{N} \sum_{b=1}^{N} \text{cov}(\hat{P}_i(\theta_a), \hat{P}_i(\theta_a)) \tag{8}
\]

\[
\text{cov}(p_i, p_j | \theta) = \frac{1}{N^2} \sum_{a=1}^{N} \sum_{b=1}^{N} \text{cov}(\hat{P}_i(\theta_a), \hat{P}_j(\theta_b)) \tag{9}
\]

The \(\text{cov}(\hat{P}_i(\theta_a), \hat{P}_j(\theta_b))\) are evaluated by applying the delta method (Kelley, 1947, pp. 524-526; Kendall and Stuart, 1969, Section 10.6) to (1). For fixed \(\theta\) (for simplicity, the notation "|\(\theta\)" is omitted from the following formula)

\[
\text{cov}(\hat{P}_i(\theta_a), \hat{P}_j(\theta_b)) = v_{ia} w_{jb} \{ t_{ia} t_{jb} \{ \text{cov}(\theta_a, \hat{\theta}_b) - \text{cov}(\hat{\theta}_b, \hat{\theta}_b) \}
\]

\[
- \text{cov}(\hat{\theta}_a, \hat{b}_j) + \text{cov}(\hat{b}_a, \hat{b}_j)) + v_{ia} t_{jb} \{ \text{cov}(\hat{a}_i, \hat{b}_b) - \text{cov}(\hat{a}_i, \hat{b}_j)) \}
\]

\[
+ v_{jb} t_{ia} \{ \text{cov}(\theta_a, \hat{a}_j) - \text{cov}(\hat{a}_i, \hat{a}_j)) + v_{ia} v_{jb} \text{cov}(\hat{a}_i, \hat{a}_j)
\]

\[
+ t_{jb} \{ \text{cov}(\hat{c}_i, \hat{b}_b) - \text{cov}(\hat{c}_i, \hat{b}_j)) / 1.7 + [v_{jb} \text{cov}(\hat{c}_i, \hat{b}_j)]
\]

\[
+ v_{ia} \text{cov}(\hat{a}_i, \hat{c}_j) / 1.7 + t_{ia} \{ \text{cov}(\theta_a, \hat{c}_j) - \text{cov}(\hat{b}_i, \hat{c}_j)) / 1.7
\]

\[
+ \text{cov}(\hat{c}_i, \hat{c}_j) / (1.7)^2 \} \tag{10}
\]
The standard errors for $\pi$ were calculated from (5) and again from (8) and (10) for each of the 45 items in the test described in Section 3. The results from the two different approaches agree to at least three significant digits for each item. The $\text{cov}(\hat{p}_1, \hat{p}_j | \theta)$ obtained from (9) and (10) were all of order $10^{-7}$ or less. This gives us increased confidence in the Lord-Wingersky sampling covariance matrix.
3. Effects of Changing Number of Items, Number of Examinees, or the Frequency Distribution of Ability

To investigate the effect of changing the number of items, the number of examinees, or the distribution of abilities on the sampling errors of parameter estimates, various sets of parameters were specified. The simplest set of parameters represents the administration of a 45-item test to 1500 examinees. The numerical values used as the true \( \theta_a \) were a spaced sample of 1500 \( \theta_a \) drawn from the ability estimates obtained by LOGIST for a regular administration of the Test of English as a Foreign Language (TOEFL). A spaced sample of fifteen items were drawn from the sixty TOEFL items whose parameters were estimated in the same run as the abilities. The estimated parameters for these fifteen items were used as the true parameters. These fifteen items were then replicated twice to get a total of 45 items, where items 16-30 and items 31-45 have the same item parameters as items 1-15. Note that various parameters were specified, but no sets of artificial data were generated for this study, since sampling variances and covariances depend only on the true parameters, not on sample observations.

To investigate the effect of increasing the number of examinees, each of 1500 \( \theta_a \) was repeated four times to represent the \( \theta_a \) of 6000 examinees. To study the effect of increasing the number of items, another 45 items were added exactly like the first 45 to create a 90-item test. For a different distribution of abilities, a rectangular distribution of 1500 \( \theta_a \) between -3 and 3 was randomly generated.
Tables 1-4 give the standard errors of the parameter estimates that would be obtained from actual data in the various situations investigated. Only the standard errors for the fifteen unique items are given in the tables of the standard errors for the item parameters. The abilities are grouped into 16 intervals between -4 and 3. Two of the intervals had no examinees. \( N \) is the number of examinees and \( n \) is the number of items. The values of both the "small" and "capital" parameters are given. The constants to convert from the small scale to the capital scale are \( \tilde{b}_0 = -0.305 \) and \( k = 0.976 \).

Figure 1 contains plots corresponding to these tables. Gaps in the curve for the \( B_i \) are due to some points out of the range of the plot. The standard error for \( C_i \) was not plotted against \( C_i \), since most of the \( C_i \) were equal, but against \( B_i - 2/A_i \) instead. \( B_i - 2/A_i \) is an indicator of the ability level at which the item response curve becomes asymptotic. The higher \( B_i - 2/A_i \), the better one should be able to estimate \( C \).

As expected, quadrupling the number of examinees halved the standard errors of the estimated item parameters; doubling the number of items, decreased the standard errors of the estimated abilities by a factor of \( \sqrt{2} \). Quadrupling the number of examinees reduces the largest standard errors for \( \hat{\Theta}_a \) sharply, but has little effect on the smaller standard errors; doubling the number of items has only a moderate or
Table 1
Standard Errors for $A_i$

<table>
<thead>
<tr>
<th>Item No.</th>
<th>$a_i$</th>
<th>$A_i$</th>
<th>n=45, N=1500</th>
<th>n=90, N=1500</th>
<th>n=45, N=6000</th>
<th>n=45, N=1500</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.99</td>
<td>0.96</td>
<td>0.234</td>
<td>0.192</td>
<td>0.117</td>
<td>0.178</td>
</tr>
<tr>
<td>2</td>
<td>0.35</td>
<td>0.34</td>
<td>0.134</td>
<td>0.131</td>
<td>0.067</td>
<td>0.072</td>
</tr>
<tr>
<td>3</td>
<td>1.38</td>
<td>1.34</td>
<td>0.318</td>
<td>0.243</td>
<td>0.159</td>
<td>0.235</td>
</tr>
<tr>
<td>4</td>
<td>0.78</td>
<td>0.76</td>
<td>0.147</td>
<td>0.126</td>
<td>0.073</td>
<td>0.099</td>
</tr>
<tr>
<td>5</td>
<td>0.4</td>
<td>0.41</td>
<td>0.100</td>
<td>0.106</td>
<td>0.050</td>
<td>0.055</td>
</tr>
<tr>
<td>6</td>
<td>0.92</td>
<td>0.90</td>
<td>0.178</td>
<td>0.145</td>
<td>0.089</td>
<td>0.120</td>
</tr>
<tr>
<td>7</td>
<td>1.06</td>
<td>1.04</td>
<td>0.209</td>
<td>0.168</td>
<td>0.104</td>
<td>0.141</td>
</tr>
<tr>
<td>8</td>
<td>1.34</td>
<td>1.31</td>
<td>0.262</td>
<td>0.205</td>
<td>0.131</td>
<td>0.180</td>
</tr>
<tr>
<td>9</td>
<td>1.50</td>
<td>1.46</td>
<td>0.317</td>
<td>0.259</td>
<td>0.158</td>
<td>0.231</td>
</tr>
<tr>
<td>10</td>
<td>0.87</td>
<td>0.85</td>
<td>0.180</td>
<td>0.151</td>
<td>0.090</td>
<td>0.117</td>
</tr>
<tr>
<td>11</td>
<td>0.62</td>
<td>0.60</td>
<td>0.142</td>
<td>0.128</td>
<td>0.071</td>
<td>0.086</td>
</tr>
<tr>
<td>12</td>
<td>1.09</td>
<td>1.06</td>
<td>0.234</td>
<td>0.197</td>
<td>0.117</td>
<td>0.153</td>
</tr>
<tr>
<td>13</td>
<td>1.39</td>
<td>1.36</td>
<td>0.311</td>
<td>0.265</td>
<td>0.156</td>
<td>0.204</td>
</tr>
<tr>
<td>14</td>
<td>1.50</td>
<td>1.46</td>
<td>0.333</td>
<td>0.283</td>
<td>0.166</td>
<td>0.209</td>
</tr>
</tbody>
</table>
Table 2

Standard Errors for $b_i$

<table>
<thead>
<tr>
<th>Item No.</th>
<th>$b_1$</th>
<th>$b_i$</th>
<th>n=45</th>
<th>n=90</th>
<th>n=45</th>
<th>n=45</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>N=1500</td>
<td>N=1500</td>
<td>N=6000</td>
<td>N=1500</td>
</tr>
<tr>
<td>1</td>
<td>-2.01</td>
<td>-1.75</td>
<td>0.516</td>
<td>0.466</td>
<td>0.258</td>
<td>0.339</td>
</tr>
<tr>
<td>2</td>
<td>-1.61</td>
<td>-1.33</td>
<td>2.544</td>
<td>2.344</td>
<td>1.272</td>
<td>1.470</td>
</tr>
<tr>
<td>3</td>
<td>-1.09</td>
<td>-0.80</td>
<td>0.353</td>
<td>0.259</td>
<td>0.177</td>
<td>0.242</td>
</tr>
<tr>
<td>4</td>
<td>-0.77</td>
<td>-0.48</td>
<td>0.257</td>
<td>0.240</td>
<td>0.128</td>
<td>0.177</td>
</tr>
<tr>
<td>5</td>
<td>-0.67</td>
<td>-0.38</td>
<td>0.965</td>
<td>0.929</td>
<td>0.483</td>
<td>0.591</td>
</tr>
<tr>
<td>6</td>
<td>-0.34</td>
<td>-0.04</td>
<td>0.191</td>
<td>0.161</td>
<td>0.095</td>
<td>0.141</td>
</tr>
<tr>
<td>7</td>
<td>-0.15</td>
<td>0.16</td>
<td>0.165</td>
<td>0.141</td>
<td>0.082</td>
<td>0.128</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
<td>0.31</td>
<td>0.143</td>
<td>0.117</td>
<td>0.071</td>
<td>0.113</td>
</tr>
<tr>
<td>9</td>
<td>0.11</td>
<td>0.42</td>
<td>0.124</td>
<td>0.096</td>
<td>0.062</td>
<td>0.096</td>
</tr>
<tr>
<td>10</td>
<td>0.26</td>
<td>0.58</td>
<td>0.110</td>
<td>0.092</td>
<td>0.055</td>
<td>0.097</td>
</tr>
<tr>
<td>11</td>
<td>0.46</td>
<td>0.79</td>
<td>0.103</td>
<td>0.101</td>
<td>0.051</td>
<td>0.098</td>
</tr>
<tr>
<td>12</td>
<td>0.57</td>
<td>0.90</td>
<td>0.178</td>
<td>0.179</td>
<td>0.089</td>
<td>0.148</td>
</tr>
<tr>
<td>13</td>
<td>0.68</td>
<td>1.01</td>
<td>0.085</td>
<td>0.086</td>
<td>0.043</td>
<td>0.086</td>
</tr>
<tr>
<td>14</td>
<td>0.90</td>
<td>1.23</td>
<td>0.082</td>
<td>0.080</td>
<td>0.041</td>
<td>0.076</td>
</tr>
<tr>
<td>15</td>
<td>1.16</td>
<td>1.50</td>
<td>0.103</td>
<td>0.089</td>
<td>0.052</td>
<td>0.077</td>
</tr>
</tbody>
</table>
### Table 3

Standard Errors for $c_{1}$

<table>
<thead>
<tr>
<th>Item No.</th>
<th>$c_{1}$</th>
<th>$c_{1}$</th>
<th>Bell-shaped distribution</th>
<th>Rectangular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$n=45$ $N=1500$</td>
<td>$n=90$ $N=1500$</td>
</tr>
<tr>
<td>1</td>
<td>0.17</td>
<td>0.17</td>
<td>0.598 0.469 0.299</td>
<td>0.316 0.409</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
<td>0.17</td>
<td>0.715 0.628 0.358</td>
<td>0.645</td>
</tr>
<tr>
<td>3</td>
<td>0.17</td>
<td>0.17</td>
<td>0.096 0.083 0.048</td>
<td>0.183 0.0</td>
</tr>
<tr>
<td>4</td>
<td>0.17</td>
<td>0.17</td>
<td>0.144 0.123 0.072</td>
<td>0.039</td>
</tr>
<tr>
<td>5</td>
<td>0.17</td>
<td>0.17</td>
<td>0.318 0.280 0.159</td>
<td>0.033</td>
</tr>
<tr>
<td>6</td>
<td>0.17</td>
<td>0.17</td>
<td>0.071 0.064 0.035</td>
<td>0.025 0.018</td>
</tr>
<tr>
<td>7</td>
<td>0.17</td>
<td>0.17</td>
<td>0.059 0.054 0.029</td>
<td>0.021</td>
</tr>
<tr>
<td>8</td>
<td>0.17</td>
<td>0.17</td>
<td>0.041 0.039 0.021</td>
<td>0.025</td>
</tr>
<tr>
<td>9</td>
<td>0.13</td>
<td>0.13</td>
<td>0.026 0.025 0.013</td>
<td>0.021</td>
</tr>
<tr>
<td>10</td>
<td>0.34</td>
<td>0.34</td>
<td>0.026 0.026 0.013</td>
<td>0.025</td>
</tr>
<tr>
<td>11</td>
<td>0.17</td>
<td>0.17</td>
<td>0.039 0.038 0.020</td>
<td>0.039</td>
</tr>
<tr>
<td>12</td>
<td>0.17</td>
<td>0.17</td>
<td>0.068 0.064 0.034</td>
<td>0.021</td>
</tr>
<tr>
<td>13</td>
<td>0.25</td>
<td>0.25</td>
<td>0.027 0.027 0.014</td>
<td>0.018</td>
</tr>
<tr>
<td>14</td>
<td>0.29</td>
<td>0.29</td>
<td>0.020 0.020 0.010</td>
<td>0.015</td>
</tr>
<tr>
<td>15</td>
<td>0.18</td>
<td>0.18</td>
<td>0.015 0.015 0.007</td>
<td>0.015</td>
</tr>
</tbody>
</table>
Table 4

Standard Errors for $\hat{\theta}_a$

<table>
<thead>
<tr>
<th>$\theta_a$</th>
<th>$\hat{\theta}_a$</th>
<th>$n=45$ N=1500</th>
<th>$n=90$ N=1500</th>
<th>$n=45$ N=6000</th>
<th>$n=45$ N=1500</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.75</td>
<td>-2.51</td>
<td>2.090</td>
<td>1.478</td>
<td>1.331</td>
<td>1.453</td>
</tr>
<tr>
<td>-2.25</td>
<td>-1.99</td>
<td>1.296</td>
<td>0.917</td>
<td>0.879</td>
<td>0.955</td>
</tr>
<tr>
<td>-1.75</td>
<td>-1.48</td>
<td>0.861</td>
<td>0.609</td>
<td>0.621</td>
<td>0.669</td>
</tr>
<tr>
<td>-1.25</td>
<td>-0.97</td>
<td>0.607</td>
<td>0.429</td>
<td>0.460</td>
<td>0.491</td>
</tr>
<tr>
<td>0.75</td>
<td>0.46</td>
<td>0.456</td>
<td>0.322</td>
<td>0.373</td>
<td>0.390</td>
</tr>
<tr>
<td>0.25</td>
<td>0.06</td>
<td>0.349</td>
<td>0.247</td>
<td>0.309</td>
<td>0.317</td>
</tr>
<tr>
<td>0.25</td>
<td>0.57</td>
<td>0.278</td>
<td>0.196</td>
<td>0.266</td>
<td>0.268</td>
</tr>
<tr>
<td>0.75</td>
<td>1.08</td>
<td>0.261</td>
<td>0.185</td>
<td>0.260</td>
<td>0.261</td>
</tr>
<tr>
<td>1.25</td>
<td>1.59</td>
<td>0.303</td>
<td>0.214</td>
<td>0.292</td>
<td>0.295</td>
</tr>
<tr>
<td>1.75</td>
<td>2.11</td>
<td>0.422</td>
<td>0.298</td>
<td>0.394</td>
<td>0.401</td>
</tr>
<tr>
<td>2.25</td>
<td>2.62</td>
<td>0.628</td>
<td>0.444</td>
<td>0.589</td>
<td>0.599</td>
</tr>
<tr>
<td>2.75</td>
<td>3.13</td>
<td>0.931</td>
<td>0.658</td>
<td>0.888</td>
<td>0.900</td>
</tr>
</tbody>
</table>
Figure 1. Comparison of the standard for $A_1$, $B_1$, $C_1$, and $\hat{\theta}_a$ for different numbers of items, different numbers of examinees and a different distribution of examinees.
small effect on the standard errors of item parameter estimates. Note that the effects discussed in the previous sentence cannot be investigated at all using the usual standard error formulas, which assume either that the item parameters are known or else that the $\theta _{a}$ are known.

The rectangular distribution of abilities definitely gives better estimates of the item parameters than the bell-shaped distribution of abilities. For $C_1$, where $B_1 = 2/A_1$ is low, the rectangular distribution gave standard errors nearly as low as the standard errors with quadruple the number of examinees.

4. Displaying Standard Errors and Sampling Covariances

In looking at tables of standard errors it is hard to see how the standard errors for $\hat{A}_1$, $\hat{B}_1$, and $\hat{C}_1$ interrelate and how the standard errors relate to the magnitude of the parameters. A plot of the three-dimensional asymptotic joint normal distribution of $\hat{A}$, $\hat{B}$, and $\hat{C}$ would be useful but difficult to read. However, projections of the contours of this distribution onto the three two-dimensional planes will give a graphical representation not only of the magnitude of the standard errors but also of the sampling correlations between the parameter estimates. The projected contours are two-dimensional ellipses. These plots are a refinement of a suggestion by Thomas Warm (personal communication, 1982).

For convenience, the subscript 1 will now be dropped. To plot the projection of the three-dimensional contour onto the $(\hat{A}, \hat{B})$-plane, only $\text{var}(\hat{A})$, $\text{var}(\hat{B})$, and $\text{cov}(\hat{A}, \hat{B})$ are needed. The exponent of
the asymptotic bivariate normal distribution of \( \hat{A} \) and \( \hat{B} \) is given by the right side of (11). The quadratic in brackets is asymptotically distributed as chi square with 2 degrees of freedom. The 95th percentile for a \( \chi^2 \) with 2 degrees of freedom is 5.99. Thus 95 percent of the time the obtained \((\hat{A}, \hat{B})\) will lie within the ellipse given by the equation

\[
5.99 = \frac{1}{1 - \rho^2} \left[ \frac{(\hat{A} - A)^2}{\text{Var}(A)} - \frac{2\rho(\hat{A} - A)(\hat{B} - B)}{\sqrt{\text{Var}(A) \text{Var}(B)}} + \frac{(\hat{B} - B)^2}{\text{Var}(B)} \right]
\]  

(11)

where

\[
\rho = \frac{\text{Cov}(\hat{A}, \hat{B})}{\sqrt{\text{Var}(A) \text{Var}(B)}}
\]

Similar equations apply for the projections onto the \((\hat{A}, \hat{C})\) and \((\hat{B}, \hat{C})\) planes. The ellipse plotted from (11) for a given \( N \) is identical to the 53-percent ellipse that would be plotted for a sample size \( N/4 \).

The following procedure was used to plot a representative set of ellipses. A hypothetical test of 60 items was created by selecting 60 items from an operational SAT mathematics test and treating these item parameter estimates as the true parameters. A standard normal distribution of 1000
abilities was generated. We then created 15 new items with all combinations of the parameters \(a = 0.5, 1.0, 1.5\); \(b = -2, -1, 0, 1, 2\); and \(c = 0.15\). Using these new items, fifteen 61-item tests were created, each containing the 60 original items and one of the new items. The sampling variance-covariance matrix for each of the fifteen 61-item tests was obtained. These matrices differ only because the 61st item differs for each matrix. Only the variances and covariances for the 61st item were used in (11) to compute the ellipses.

The plots were made for an \(N\) of 16,000 to avoid confusing overlap of the ellipses. These ellipses are also the 53% confidence ellipses for an \(N\) of 4000. The left and bottom axes are labeled with the "small" scale, the right and top axes are labeled with the "capital" scale. The standard errors used are for parameter estimates on the capital scale. The transformation parameters to transform from the small to the capital scale are \(k = 1.336\). The center of the ellipse is marked by a "+".

Figure 2 shows the ellipses on the \((A, B)\) plane. The plot shows that the standard error of \(\hat{A}\) increases with \(A\). The standard error of \(\hat{B}\) increases as \(B\) approaches the extremes. The sampling correlation between \(\hat{A}\) and \(\hat{B}\) is moderately or strongly positive for easy items and moderately or strongly negative for hard items.
Figure 2. Projections onto the $(A,B)$-plane of the 95% ellipses for an $N$ of 16,000.
Figure 3 shows the projections onto the \((B,C)\) plane. At each value of \(B\) there are three ellipses, which are concentric because \(c = C = .15\) for all items. The longest ellipse along the \(C\) axis is for \(a = .5\), the middle ellipse is for \(a = 1.0\), and the shortest is for \(a = 1.5\). The other triples of ellipses are similarly ordered on \(a\). The standard error of \(\hat{C}\) is large for easy items and moderately small for difficult items; the standard error of \(\hat{C}\) decreases as \(a\) increases. As \(a\) decreases, the sampling correlation between \(B\) and \(\hat{C}\) becomes strongly positive except for hard items where \(\hat{C}\) is well determined.

Figure 4 shows the projections onto the \((A,C)\) plane. There are five concentric ellipses for each value of \(A\). The ellipse with the longest \(c\) axis is for \(b = -2.0\), the ellipse with the shortest \(c\) axis is for \(b = 2.0\). Again \(\hat{C}\) has large standard errors for easy items and for items with low \(a\)'s. For hard items the sampling correlation between \(A\) and \(\hat{C}\) is positive and sometimes high; for easy items, the correlation is negative.

4. Standard Errors for Two Tests with Common Items

Suppose that each of two tests measuring the same ability is administered to a different group of examinees. We want to use item response theory either to put the items for both tests into a common item pool or to equate the two tests. For either purpose it is necessary that all the estimated parameters be on the same scale.
Figure 3. Projections onto the $(B,C)$-plane of the 95% ellipses for an $N$ of 16,000.
Figure 4. Projections onto the \((A, C)\) -plane of the 95\% ellipses for an \(N\) of 16,000.
Unless equivalent groups of examinees are used, methods for doing this usually require a subset of items that are common to both tests. The unique items are the items in each test that are not common to the other test. The item parameters for each test can then either be estimated separately in two calibration runs or together in one calibration run. If the parameters are estimated in two separate runs, there are two different parameter estimates for each common item. These should be the same except for sampling error and the arbitrary origin and unit of measurement of the ability scale. There are several methods for determining the linear transformation necessary to transform the item parameter estimates for both tests to the same scale. These methods will not be described here (see Stocking and Lord, 1983). However, if all of the items for both tests are calibrated in one run, called a concurrent calibration, the parameters for both tests are automatically put on the same scale and no linear transformation is necessary. This concurrent procedure is most efficient; it provides smaller standard errors and involves fewer assumptions than other procedures. The concurrent procedure is the procedure studied here.

One question that arises when applying the common item method for putting the parameters for both tests on a common scale is: How many common items are necessary? Vale, Maurelli, Gialluca, Weiss, and Ree (1981) investigated this problem using simulated data with 5, 15, and 25 common items and three different shapes of the common item section test information curve: peaked, normal, and rectangular. They also investigated many other linking methods. For the common item method, they assumed that one already had good estimates of the parameters for the common items and required that one have enough common and unique items to get good estimates of the abilities. They
used two estimates of the abilities, one obtained from the common items, the
other from the unique items to determine the transformation to put the unique
items onto the common scale. They found that 15 to 25 items were necessary and
that the common item sections with a rectangular or normal information function
were better than those with a peaked information function.

Another study to determine the number of common items necessary was done
by McKinley and Reckase (1981). They compared the concurrent method and
several other methods for obtaining the linear transformations using the
two sets of item parameter estimates for the common items. A large set of
items using real data from a multidimensional achievement test covering seven
subareas was calibrated in one calibration run and these parameter estimates
were used as the criterion for determining how well the other linking
procedures put the parameter estimates for subsets of these items on a common
scale. A chain of three links was created, that is, test A was linked to
test B through one set of common items, test B to test C through another
set of common items, and test C to test D through a third set. Five sample
sizes ranging from 100 examinees to 2000 examinees were used. All four tests
were then calibrated in one run for the concurrent method for each sample. The
linking was done with 5, 15 and 25 common items. Each individual test was 50
items long including the common items. McKinley and Reckase concluded
that 5 items were not adequate, 25 items were better than 15, but 15 were
adequate for linking with the concurrent method.

Given the sampling variance-covariance matrix for all parameter estimates
in our single concurrent run when all parameters are treated as unknown, we
can investigate what effect the number of common items has on the sampling standard errors of the unique items in both tests. Note that this problem cannot be investigated at all with the limited sampling-error formulas that assume that either item or ability parameters are known.

Numerical Procedures

Suppose test 1 has a section of unique items labeled $V_4$, and test 2 has a section of unique items labeled $Z_5$. Both tests have the same set of common items labeled $C_0$. One group of examinees, group X, took test 1, another group of examinees, group Y, took test 2. The information matrix $I_{pq}$, which must be inverted to get the variance-covariance matrix, has the following structure (Lord and Wingersky, 1983):

\[
I_{pq} = \begin{bmatrix}
S_{11} & 0 & 0 & F_{11} & 0 \\
0 & S_{22} & 0 & F_{21} & F_{22} \\
0 & 0 & S_{33} & 0 & F_{32} \\
F_{11} & F_{21} & 0 & T_{11} & 0 \\
0 & F_{22} & F_{32} & 0 & T_{22}
\end{bmatrix}
\]
The $S$ submatrices ($S_{11}$ for the V4 items; $S_{22}$ for the common items; $S_{33}$ for the Z5 items) contain $3 \times 3$ Fisher information matrices for $a_i$, $b_i$, $c_i$ on the diagonal. The $T$ submatrices are the diagonal information matrices for the examinees: $T_{11}$ for the examinees that took test 1; $T_{22}$ for the examinees that took test 2. The $P$ submatrices contain the vectors $f_{ia}$, each of which is the $3 \times 1$ Fisher information vector for item $i$ and examinee $a$. Note that for Group $Y$, this is 0 for the V4 items; for Group $X$, this is 0 for Z5.

The matrix $N_{pq}$ is inverted by grouping the abilities for group $X$ into sixteen groups and by grouping the abilities for group $Y$ into another set of sixteen groups. Then the formulas for inverting a partitioned matrix using the method described in Lord and Wingersky (1983) are successively applied.

**Data and Results**

To study the effect of the number of common items on the standard errors of the parameter estimates for the unique items, we selected two 60-item SAT Mathematics tests with an additional 25-item common-item section. The 60 unique items in the first test will be referred to as V4 and the 60 unique items in the second test will be referred to as Z5. Estimates of all of the parameters were obtained in one concurrent LOGIST run. These estimates were treated as true parameter values in computing the standard errors for all 145 items.
We then doubled the length of the common item section by simply replicating the parameters for the 25 common items. Surprisingly, the standard errors for the 120 unique items in V4 and Z5 computed with 50 common items agreed with the standard errors computed with only 25 common items to two decimal places. If doubling the number of common items makes so little difference, what is the effect of halving the number of common items? Or at the extreme, reducing the number of common items to 2?

This is really not as absurd as it sounds. Providing the common items are not part of the test score, other than improving the estimates of the abilities, the function of the common items is to put the parameters for the two sets of unique items on the same metric. If the model holds, only a linear transformation is required to convert the parameters from one scale to another. Only 2 parameters are necessary to determine this linear transformation. With 2 common items we are estimating four parameters that affect the scale, the two \( a \)'s influence the scale unit and the two \( b \)'s influence both the scale unit and origin. The two \( c \)'s are not affected by the scale. Consequently with 2 items we actually have two more parameters than absolutely necessary. However, if the 2 common items have parameter estimates with large standard errors, the scale will be less well determined than if the estimates have small standard errors.

To study the effect of two common items on the standard errors of the unique items, we selected 2 "good" items and 2 "bad" items from the 25 common items. The item parameters and their standard errors for the 2 "good" items were
The item parameters and their standard errors for the 2 "bad" common items were

<table>
<thead>
<tr>
<th>a</th>
<th>SE(A)</th>
<th>b</th>
<th>SE(B)</th>
<th>c</th>
<th>SE(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>.98</td>
<td>.09</td>
<td>-.10</td>
<td>.02</td>
<td>.06</td>
<td>.02</td>
</tr>
<tr>
<td>.96</td>
<td>.10</td>
<td>.21</td>
<td>.02</td>
<td>.15</td>
<td>.02</td>
</tr>
</tbody>
</table>

These standard errors were computed for the situation where all 25 common items are included in the parameter estimation run.

We then obtained the variance-covariance matrix for the V4 and V5 items when only the 2 good common items are included in the estimation run and also the variance-covariance matrix when only the 2 bad common items are used.

The constants to transform from the small scale to the capital scale are \( \hat{\alpha} = -0.261 \) and \( \hat{\beta} = 1.914 \). Only V4 and V5 items were used to compute \( \hat{\alpha} \) and \( \hat{\beta} \), so that the same transformation would apply to all four variance-covariance matrices.

Table 5 gives the medians, and the bottom and top quartiles of the standard errors for \( \hat{A} \), \( \hat{B} \), and \( \hat{C} \), for the Z4 and Z5 unique items computed for four different situations: using 50 common items, using 25 common items, using 2 good common items, and using 2 bad common items. Using 2 good common items gives smaller standard errors for the unique items than using 2 bad common items. The standard errors using the 2 good items...
Table 5
Comparison of the Standard Errors of Estimated Item Parameters across the Four Sets of Common Items

<table>
<thead>
<tr>
<th></th>
<th>50 Common Items</th>
<th>25 Common Items</th>
<th>2 Good Common Items</th>
<th>2 Bad Common Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Errors for A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Quartile</td>
<td>0.114</td>
<td>0.115</td>
<td>0.123</td>
<td>0.131</td>
</tr>
<tr>
<td>Median</td>
<td>0.140</td>
<td>0.141</td>
<td>0.151</td>
<td>0.163</td>
</tr>
<tr>
<td>Third Quartile</td>
<td>0.224</td>
<td>0.226</td>
<td>0.236</td>
<td>0.243</td>
</tr>
<tr>
<td>Standard Errors for B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Quartile</td>
<td>0.029</td>
<td>0.030</td>
<td>0.034</td>
<td>0.041</td>
</tr>
<tr>
<td>Median</td>
<td>0.042</td>
<td>0.042</td>
<td>0.048</td>
<td>0.056</td>
</tr>
<tr>
<td>Third Quartile</td>
<td>0.066</td>
<td>0.067</td>
<td>0.072</td>
<td>0.076</td>
</tr>
<tr>
<td>Standard Errors for C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Quartile</td>
<td>0.013</td>
<td>0.013</td>
<td>0.013</td>
<td>0.013</td>
</tr>
<tr>
<td>Median</td>
<td>0.027</td>
<td>0.027</td>
<td>0.028</td>
<td>0.027</td>
</tr>
<tr>
<td>Third Quartile</td>
<td>0.055</td>
<td>0.055</td>
<td>0.058</td>
<td>0.056</td>
</tr>
</tbody>
</table>
are not much larger than the standard errors using 25 common items. Even reliance on just 2 bad common items gives surprisingly good results. Since the purpose of the common items is to determine the scale, it is not surprising that the number of common items has a negligible effect on the standard error of $e$, since $c$ is independent of the ability scale.

Table 6 gives the standard errors for the abilities computed with the four different sets of common items. Not surprisingly, if we increase the number of common items to 50 we reduce the standard error of the abilities, although not uniformly as shown by the ratio column. The standard error for the abilities at -2 were lower when computed using the two bad common items, which were easy items, than when computed using the two good common items.

Even though there is little difference between the standard errors when there are 2 common items and when there are 25 common items, the parameter estimates for the V4 and Z5 items will not have been adequately put on the same scale if all of the parameter estimates for V4 items err in one direction and all of the parameter estimates for Z5 items err in the opposite direction. Is this what will happen in practice? To determine how well an anchor-test of only 2 common items puts tests V4 and Z5 on the same scale, we reestimated the parameters twice, once in a LOGIST run with the items for Z5 and V4 and the two "good" common items, the other in a LOGIST run with the items for Z5 and V4 and the two "bad" common items.

The estimated parameters for Z5 and V4 computed with the 25 common items will be used as the criterion for evaluating the calibrations.
Table 6
Comparison of the Standard Errors of Estimated Abilities across the Four Sets of Common Items

<table>
<thead>
<tr>
<th>θ</th>
<th>50 Common Items</th>
<th>25 Common Items</th>
<th>2 Good Common Items</th>
<th>2 Bad Common Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ</td>
<td>S.E.</td>
<td>S.E.</td>
<td>Ratio</td>
<td>S.E.</td>
</tr>
<tr>
<td>2.00</td>
<td>1.18</td>
<td>0.097</td>
<td>0.109</td>
<td>0.894</td>
</tr>
<tr>
<td>1.00</td>
<td>0.66</td>
<td>0.089</td>
<td>0.102</td>
<td>0.870</td>
</tr>
<tr>
<td>0.0</td>
<td>0.14</td>
<td>0.100</td>
<td>0.115</td>
<td>0.874</td>
</tr>
<tr>
<td>-1.00</td>
<td>-0.39</td>
<td>0.129</td>
<td>0.145</td>
<td>0.892</td>
</tr>
<tr>
<td>-2.00</td>
<td>-0.91</td>
<td>0.221</td>
<td>0.248</td>
<td>0.891</td>
</tr>
</tbody>
</table>
with 2 common items. The 2 good common items did fairly well at putting the parameters on this scale. The 2 bad items did not do so well.

The top plot in Figure 5 compares the \( b \)'s for the 60 unique \( V4 \) items estimated with 2 good items with the \( b \)'s estimated with 25 common items. Similarly, the bottom plot compares the \( b \)'s for the unique \( Z5 \) items. If the parameters were on the same metric the \( b \)'s in both plots should fall on a 45° line. The difference from the 45° line is hard to distinguish. The two points for \( Z5 \) that are far away from the 45° line had the \( c \)'s fixed by LOGIST at the common \( c \) value in one calibration but not in the other.

Figure 6 shows the plots for the \( a \)'s for \( V4 \) and \( Z5 \) respectively. Here it definitely looks as if the \( a \)'s are not on the same scale. The \( a \)'s for the \( V4 \) items have a slope greater than 45°.

Figure 7 compares the \( b \)'s estimated with the 2 bad common items with the \( b \)'s estimated with 25 common items. Here the points for the \( V4 \) items are above the 45° line, and points for the \( Z5 \) items are below the line. The plots comparing the \( a \)'s in Figure 8 confirm that the 2 bad common items do not put the parameters for \( Z5 \) and \( V4 \) on the same metric. As suspected, with the 2 bad items the parameters for one set of the unique items err in one direction and for the other set, in the opposite direction.

The reason for putting \( Z5 \) and \( V4 \) on the same scale was to equate \( Z5 \) to \( V4 \) using true-score equating. What effect does using only 2 common items to put the two forms on the same scale have on the true-score equating
Figure 5. Comparison of the $b$'s estimated with 2 good common items and the $b$'s estimated with 25 common items, separately for $V4$ and $Z5$. 
Figure 6. Comparison of the $a^\prime$'s estimated with 2 good common items and the $a^\prime$'s estimated with 25 common items, separately for V4 and Z5.
Figure 7. Comparison of the b's estimated with 2 bad common items and the b's estimated with 25 common items, separately for V4 and Z5.
Figure 8. Comparison of the $a$'s estimated with 2 bad common items and the $a$'s estimated with 25 common items, separately for V4 and Z5.
between the two forms? Figure 9 shows three true-score equating lines: the solid line is the equating line found when the parameters are estimated using 25 common items, the dotted line is the equating line found when the parameters are estimated using just the 2 good common items, the dashed line is found when the parameters are estimated using just the 2 bad common items. For this equating, true scores on form 25 are first equated to true scores on V4. Then the true scores on V4, are converted to scaled scores between 100 and 800 by a linear transformation. Using the equating line with the 25 items as a criterion, the equating using 2 bad common items is worse than the equating using 2 good common items. The equating using the 2 good common items is close to the equating with 25 common items; the maximum scaled score difference is 8 points.

All of these results assume that the item parameters estimated using 25 common items are on the same scale. This analysis should be repeated in a situation where one knows that all of the parameters used as a criterion are on a common scale. From the results so far, it appears that good linking may be obtained with as few as five common items or less. However, these results only apply when the item parameters for the two forms are put on a common scale by estimating all of them in one calibration run. These results do not apply when the two tests are calibrated in two separate runs and the parameters are put on a common scale using some linear transformation determined from the common items.
Figure 9. Comparisons of the three true-score equatings of test 25 to test V4: using 25 common items, using 2 good common items, and using 2 bad common items.
The conclusion that good linking may be obtained with as few as five commonly used items is more optimistic than the conclusions reached by Vale et al. (1981) and by McKinley and Reckase (1981). Our differences with Vale et al. may be due to the facts that 1) their scaling was based on estimated θ's, and 2) they used three estimation runs instead of one concurrent run. Our differences with McKinley and Reckase are probably due to the facts that in their study 1) the responses of some examinees to some items (as we understand it) often appeared twice in the same concurrent LOGIST run, violating the assumption of local independence; and, more importantly, 2) they pooled the Iowa Tests of Educational Development covering seven different achievement areas, and analyzed the resulting multidimensional pool of items as if it were unidimensional.

**Summary**

The asymptotic sampling variance-covariance matrix of maximum likelihood estimators when both abilities and item parameters are unknown was used to study several problems in item response theory, such as the extent to which more items, more examinees, or a different distribution of abilities will provide better estimates of parameters. It was found for the values of n and N studied that that the standard error of \( \hat{\theta} \) varies inversely as \( \sqrt{n} \), but is only moderately affected by changes in N; the standard error of the estimated item parameters varies inversely as \( \sqrt{N} \), but is only slightly affected by changes in n.
A rectangular distribution of abilities gives smaller standard errors for the item parameters than doubling the number of items. In fact, for low A's, also for C's, for items with $B - 2/A$ less than -1, the standard errors computed with a rectangular distribution of ability were nearly as low as the standard errors computed with a bell-shaped distribution and quadruple the number of people.

With the variance-covariance matrix computed when all parameters are treated as unknown, one can study the effect of the number of common items on the standard errors of the unique items when each of two tests containing common items is administered to a different group of examinees and the parameters for both tests are calibrated in one LOGIST run. This problem cannot be dealt with at all by previously available sampling error formulas. The number of common items has little effect on the standard errors of the parameters for the unique items. The standard errors indicate that as few as 2 items may be sufficient providing the parameter estimates for these two items are well determined. However when two tests were actually calibrated in one LOGIST run using 2 common items that had parameter estimates with low standard errors, the parameters were not quite on the same scale as the parameters estimated with 25 common items. The $b$'s were very close to the same scale but the $a$'s for one of the tests were on a slightly different scale. Although 2 items are not quite enough, adequate linking may be possible with as few as five items.
References


Warm, T. A. Personal communication, 1981.

DISTRIBUTION LIST

Navy

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

1 Dr. Arthur Bachrach
Environmental Stress Program Center
Naval Medical Research Institute
Bethesda, MD 20014

1 Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152

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

1 Lt. Alexander Bory
Applied Psychology
Measurement Division
NAMRL
NAS Pensacola, FL 32508

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

1 Dr. Robert Carroll
NAVOP 115
Washington, DC 20370

1 Chief of Naval Education and Training Liaison Office
Air Force Human Resource Laboratory
Flying Training Division
Williams Air Force Base, AZ 85224

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

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

1 Dr. Tom Duffy
Navy Personnel R&D Center
San Diego, CA 92152

1 Mr. Mike Durmeyer
Instructional Program Development
Building 90
NET-PDCD
Great Lakes NTC, IL 60088

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

1 Dr. Pat Federico
Code P13
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Cathy Fernandes
Navy Personnel R&D Center
San Diego, CA 92152

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

1 Dr. Jim Hollan
Code 14
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Ed Hutchins
Navy Personnel R&D Center
San Diego, CA 92152
1 Dr. Norman J. Kerr  
Chief of Naval Technical Training  
Naval Air Station Memphis (75)  
Millington, TN 38054

1 Dr. Peter Kincaid  
Training Analysis & Evaluation Group  
Department of the Navy  
Orlando, FL 32813

1 Dr. R. W. King  
Director, Naval Education and Training Program  
Naval Training Center, Bldg. 90  
Great Lakes, IL 60088

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

1 Dr. William L. Maloy (02)  
Chief of Naval Education and Training  
Naval Air Station  
Pensacola, FL 32508

1 Dr. Kneale Marshall  
Chairman, Operations Research Dept.  
Naval Post Graduate School  
Monterey, CA 93940

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

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

1 Mr. William Nordbrock  
1032 Fairlawn Avenue  
Libertyville, IL 60048

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

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

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

1 Special Asst. for Education and Training (OP-01E)  
Room 2705, Arlington Annex  
Washington, DC 20370

1 LT Frank C. Petho, MSC, USN  
CNET (N-432)  
NAS  
Pensacola, FL 32508

1 Dr. Bernard Rimland (01C)  
Navy Personnel R & D Center  
San Diego, CA 92152

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

1 Dr. Worth Scanland, Director  
CNET (N-5)  
NAS  
Pensacola, FL 32508

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

1 Dr. Alfred F. Smode, Director  
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
<table>
<thead>
<tr>
<th>Name</th>
<th>Position/Role</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr. Frederick Steinheiser</td>
<td>Director, Office of Manpower Utilization</td>
<td>CNO - OP115, Navy Annex, Arlington, VA 20370</td>
</tr>
<tr>
<td>Mr. Brad Symson</td>
<td>Headquarter, U.S. Marine Corps</td>
<td>Naval Personnel R &amp; D Center, San Diego, CA 92152</td>
</tr>
<tr>
<td>Dr. Frank Vicino</td>
<td>Special Assistant for Marine Corps Matters</td>
<td>Naval Personnel R &amp; D Center, San Diego, CA 92152</td>
</tr>
<tr>
<td>Dr. Edward Wegman</td>
<td>Director, Office of Manpower Utilization</td>
<td>Office of Naval Research, 800 North Quincy Street, Arlington, VA 22217</td>
</tr>
<tr>
<td>Dr. Ronald Weitzman</td>
<td>Special Assistant for Marine Corps Matters</td>
<td>Department of Administrative Services, U.S. Naval Postgraduate School, Monterey, CA 93940</td>
</tr>
<tr>
<td>Dr. Douglas Wetzel</td>
<td>Director, Office of Manpower Utilization</td>
<td>Navy Personnel R &amp; D Center, San Diego, CA 92152</td>
</tr>
<tr>
<td>Dr. Martin F. Wiskoff</td>
<td>Special Assistant for Marine Corps Matters</td>
<td>Navy Personnel R &amp; D Center, San Diego, CA 92152</td>
</tr>
<tr>
<td>Mr. John H. Wolfe</td>
<td>Special Assistant for Marine Corps Matters</td>
<td>Navy Personnel R &amp; D Center, San Diego, CA 92152</td>
</tr>
<tr>
<td>Dr. Wallace Wulfek, III</td>
<td>Special Assistant for Marine Corps Matters</td>
<td>Navy Personnel R &amp; D Center, San Diego, CA 92152</td>
</tr>
<tr>
<td>Dr. H. William Greenup</td>
<td>Director, Office of Manpower Utilization</td>
<td>Education Advisor (EO31), Education Center, MCDEC, Quantico, VA 22134</td>
</tr>
<tr>
<td>Dr. A. L. Slafkosky</td>
<td>Technical Director, U.S. Army Research Institute for the Behavioral and Social Sciences</td>
<td>5001 Eisenhower Avenue, Alexandria, VA 22333</td>
</tr>
<tr>
<td>Major Frank Yohannan, USMC</td>
<td>Technical Director, U.S. Army Research Institute for the Behavioral and Social Sciences</td>
<td>5001 Eisenhower Avenue, Alexandria, VA 22333</td>
</tr>
<tr>
<td>Dr. Kent Eaton</td>
<td>Technical Director, U.S. Army Research Institute for the Behavioral and Social Sciences</td>
<td>5001 Eisenhower Avenue, Alexandria, VA 22333</td>
</tr>
</tbody>
</table>
1 Dr. Beatrice J. Farr
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

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

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

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

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

1 Dr. Joseph Psotka
ATTN: PERI-IC
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

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

1 Dr. Joyce Shields
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Hilda Wing
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Wisher
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Laboratory
AFHRL/HPD
Brooks Air Force Base, TX 78235

1 Technical Documents Center
Air Force Human Resources Laboratory
WPAFB, OH 45433

1 U.S. Air Force Office of
Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

1 Air University Library
AUL/LSE 76/443
Maxwell AFB, AL 36112

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks Air Force Base, TX 78235

1 Mr. Raymond E. Christal
AFHRL/NOE
Brooks AFB, TX 78235

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

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

1 Dr. T. N. Longridge
AFHRL/OTE
Williams AFB, AZ 85224

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

1 Dr. Malcolm Ree
AFHRL/HP
Brooks Air Force Base, TX 78235

1 LT Tallarigo
3700 TCHTW/TGIBR
Sheppard AFB, TX 76311

1 Dr. Joseph Yasatuke
AFHRL/LRT
Lowry AFB, CO 80230

Department of Defense

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

1 Dr. Craig I. Fields
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

1 Dr. William Graham
Testing Directorate
NEPCON/HEPCT-P
Ft. Sheridan, IL 60037

Department of Defense

1 Mr. Jerry Lehnus
HQ NEPCON
Attn: HEPCT-P
Ft. Sheridan, IL 60037

1 Military Assistant for Training and Personnel Technology
Office of the Under Secretary of Defense for Research and Engineering
Room 3DL29, The Pentagon
Washington, DC 20334

1 Dr. Wayne Sellman
Office of the Assistant Secretary of Defense (HRA & L)
7B269 The Pentagon
Washington, DC 20301

1 Major Jac Lint
DARPA
1400 Wilson Blvd.
Arlington, VA 22209

Civilian Agencies

1 Dr. Patricia A. Butler
NIE-BRN Bldg., Stop #7
1200 19th Street, NW
Washington, DC 20208

1 Dr. Susan Chipman
Learning and Development
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Arthur Nelmed
724 Brown
U.S. Department of Education
Washington, DC 20208

1 Dr. Andrew R. Molnar
Office of Scientific and Engineering Personnel and Education
National Science Foundation
Washington, DC 20550
Civilian Agencies

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

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

1. Dr. Frank Withrow
   U.S. Office of Education
   400 Maryland Avenue, SW
   Washington, DC 20202

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

Private Sector

1. Dr. James Algina
   University of Florida
   Gainesville, FL 32611

1. Dr. Patricia Baggett
   Department of Psychology
   University of Colorado
   Boulder, CO 80309

1. Dr. Isaac Bejar
   Educational Testing Service
   Princeton, NJ 08541

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

Private Sector

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

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

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

1. Dr. Ernest R. Cadotte
   307 Stokely
   University of Tennessee
   Knoxville, TN 37916

1. Dr. Pat Carpenter
   Department of Psychology
   Carnegie-Mellon University
   Pittsburgh, PA 15213

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

1. Dr. Norman Cliff
   Department of Psychology
   University of Southern California
   University Park
   Los Angeles, CA 90007

1. Dr. Allan M. Collins
   Bolt, Beranek, and Newman, Inc.
   50 Moulton Street
   Cambridge, MA 02138

1. Dr. Lynn A. Cooper
   LRDC
   University of Pittsburgh
   3939 O'Hara Street
   Pittsburgh, PA 15213
1. Dr. Hans Crombag  
   Education Research Center  
   University of Leyden  
   Hoervuvelaan 2  
   2334 EN Leyden  
   THE NETHERLANDS

1. Dr. Dattpradad Divgi  
   Syracuse University  
   Department of Psychology  
   Syracuse, NY 3210

1. Dr. Susan Emberton  
   Psychology Department  
   University of Kansas  
   Lawrence, KS 66045

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

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

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

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

1. Dr. Dexter Fletcher  
   WICAT Research Institute  
   1875 S. State Street  
   Orebro, UT 22333

1. Dr. John R. Frederiksen  
   Bolt, Beranek, and Newman  
   50 Moulton Street  
   Cambridge, MA 02138

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

1. Dr. Robert Glaser  
   LRDC  
   University of Pittsburgh  
   3939 O'Hara Street  
   Pittsburgh, PA 15213

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

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

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

1. Dr. Lloyd Humphreys  
   Department of Psychology  
   University of Illinois  
   Champaign, IL 61820

1. Dr. Jack Hunter  
   2122 Coolidge Street  
   Lansing, MI 48906

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

1. Dr. Douglas H. Jones  
   10 Trafalgar Court  
   Lawrenceville, NJ 08648
### Private Sector

1. **Prof. John A. Keats**  
   Department of Psychology  
   University of Newcastle  
   Newcastle, New South Wales 2308  
   AUSTRALIA

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

1. **Dr. Pat Langley**  
   The Robotics Institute  
   Carnegie-Mellon University  
   Pittsburgh, PA 15213

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

1. **Dr. Michael Levine**  
   Department of Educational Psychology  
   University of Illinois  
   Champaign, IL 61801

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

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

1. **Mr. Phillip Livingston**  
   Systems and Applied Sciences Corporation  
   66111 Kenilworth Avenue  
   Riverdale, MD 20840

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

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

1. **Dr. James Lumsden**  
   Department of Psychology  
   University of Western Australia  
   Nedlands, Western Australia 6009  
   AUSTRALIA

1. **Dr. Gary Marco**  
   Stop 31-E  
   Educational Testing Service  
   Princeton, NJ 08541

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

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

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

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

1. **Dr. Allen Munro**  
   Behavioral Technology Laboratories  
   1845 Elena Avenue, Fourth Floor  
   Redondo Beach, CA 90277
<table>
<thead>
<tr>
<th>Name</th>
<th>Institution and Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr. Alan Nicewander</td>
<td>University of Oklahoma, Department of Psychology, Oklahoma City, OK 73069</td>
</tr>
<tr>
<td>Dr. Donald A. Norman</td>
<td>Cognitive Science, C-015, University of California, San Diego, La Jolla, CA 92093</td>
</tr>
<tr>
<td>Dr. Melvin R. Novick</td>
<td>356 Lindquist Center for Measurement, University of Iowa, Iowa City, IA 52242</td>
</tr>
<tr>
<td>Dr. James Olson</td>
<td>WICAT, Inc., 1875 S. State Street, Orem, UT 84057</td>
</tr>
<tr>
<td>Dr. Wayne M. Patience</td>
<td>American Council on Education, CED Testing Service, Suite 20, One Dupont Circle, NW, Washington, DC 20036</td>
</tr>
<tr>
<td>Dr. James A. Paulson</td>
<td>Portland State University, P.O. Box 751, Portland, OR 97207</td>
</tr>
<tr>
<td>Dr. James W. Pellegrino</td>
<td>University of California, Santa Barbara, Department of Psychology, Santa Barbara, CA 93106</td>
</tr>
<tr>
<td>Dr. Mark D. Reckase</td>
<td>ACT, P.O. Box 168, Iowa City, IA 52243</td>
</tr>
<tr>
<td>Dr. Laureen Resnick</td>
<td>LRDC, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15261</td>
</tr>
<tr>
<td>Dr. Thomas Reynolds</td>
<td>University of Texas, Dallas, Marketing Department, P.O. Box 688, Richardson, TX 75080</td>
</tr>
<tr>
<td>Dr. Andrew Rose</td>
<td>American Institutes for Research, 1055 Thomas Jefferson St., NW, Washington, DC 20007</td>
</tr>
<tr>
<td>Dr. Ernst Z. Rothkopf</td>
<td>Bell Laboratories, Murray Hill, NJ 07974</td>
</tr>
<tr>
<td>Dr. Lawrence Rudner</td>
<td>403 Elm Avenue, Takoma Park, MD 20012</td>
</tr>
<tr>
<td>Dr. J. Ryan</td>
<td>Department of Education, University of South Carolina, Columbia, SC 29208</td>
</tr>
<tr>
<td>Dr. Fumiko Samejima</td>
<td>Department of Psychology, University of Tennessee, Knoxville, TN 37916</td>
</tr>
<tr>
<td>Dr. Walter Schneider</td>
<td>Psychology Department, 603 E. Daniel, Champaign, IL 61820</td>
</tr>
<tr>
<td>Dr. Lowell Schoer</td>
<td>Psychological and Quantitative Foundations, College of Education, University of Iowa, Iowa City, IA 52242</td>
</tr>
<tr>
<td>Dr. Robert J. Seidel</td>
<td>Instructional Technology Group, HUNRRD, 300 N. Washington Street, Alexandria, VA 22314</td>
</tr>
</tbody>
</table>
Private Sector

1 Dr. Kazuo Shigematsu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 980
JAPAN

1 Dr. Edwin Shirkey
Department of Psychology
University of Central Florida
Orlando, FL 32816

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

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

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

1 Dr. Kathryn T. Spoehr
Psychology Department
Brown University
Providence, RI 02912

1 Dr. Robert Sternberg
Department of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

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

Private Sector

1 Dr. William Stout
University of Illinois
Department of Mathematics
Urbana, IL 61801

1 Dr. Patrick Suppes
Institute for Mathematical Studies
Stanford University
Stanford, CA 94305

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

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

1 Dr. Maurice Tatsuoka
220 Education Building
1310 S. Sixth Street
Champaign, IL 61820

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

1 Dr. Douglas Towne
University of Southern California
Behavioral Technology Labs
1845 S. Elena Avenue
Redondo Beach, CA 90277

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

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

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

1 Dr. Kurt Van Lehn
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304

1 Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

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

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

1 Dr. Phyllis Weaver
2979 Alexis Drive
Palo Alto, CA 94304

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

1 Dr. Keith T. Wescoat
Perceptronics, Inc.
545 Middlefield Road
Suite 140
Menlo Park, CA 94025

Private Sector

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

1 Dr. Wolfgang Wildgrube
Streitkraefteamt
Box 20 50 03
D-5300 Bonn 2
WEST GERMANY

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

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