RR-88-38-ONR

4

SOME CONSEQUENCES OF THE UNCERTAINTY IN IRT LINKING PROCEDURES

Kathleen M. Sheehan

and

Robert J. Mislevy

This research was sponsored in part by the
Cognitive Science Program
Cognitive and Neural Sciences Division
Office of Naval Research, under
Contract No. N00014-88-K-0304
R&T 4421552

Robert J. Mislevy, Principal Investigator

Educational Testing Service
Princeton, New Jersey

July 1988

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a REPORT SECURITY CLASSIFICATION: Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): RR-88-38-ONR	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION: Educational Testing Service; 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION: Cognitive Science Program, Office of Naval Research (Code 1142CS), 800 North Quincy Street
6c ADDRESS (City, State, and ZIP Code): Princeton, NJ 08541	7b ADDRESS (City, State, and ZIP Code): Arlington, VA 22217-5000
8a NAME OF FUNDING/SPONSORING ORGANIZATION; 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-88-K-0304
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 61153N; PROJECT NO RR04204; TASK NO RR04204-01; WORK UNIT ACCESSION NO R&T4421552

11 TITLE (Include Security Classification)

Some Consequences of the Uncertainty in IRT Linking Procedures (Unclassified)

12 PERSONAL AUTHOR(S)

Kathleen M. Sheehan and Robert J. Mislevy

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM ______ TO ______	July 1988	30

16 SUPPLEMENTARY NOTATION

17 COSATI CODES

FIELD	GROUP	SUB-GROUP
05	10	

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

Item Response Theory; Linking Transformation; The Stocking-Lord Linking Procedure

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

In many practical applications of item response theory, the parameters of overlapping subsets of test items are estimated from different samples of examinees. A linking procedure is then employed to place the resulting item parameter estimates onto a common scale. It is standard practice to ignore the uncertainty associated with the linking step when drawing inferences that involve items from different subsets, a situation that arises, for example, in the measurement of change. This paper outlines how the uncertainty can be accounted for, and exemplifies the ideas with a jackknife approximation for the Stocking-Lord linking procedure. Examples from the National Assessment of Educational Progress suggest that the resulting uncertainty will usually be negligible for inferences about individuals, but can constitute a major source of estimation error in aggregate statistics such as changes in group means.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL: Dr. James Lester	22b TELEPHONE (Include Area Code): 202-696-4503; 22c OFFICE SYMBOL: ONR 1142CS

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

S/N 0102-LF-014-6603

Unclassified

Some Consequences of the Uncertainty in

IRT Linking Procedures

Kathleen M. Sheehan and Robert J. Mislevy

Educational Testing Service

July, 1988

The issue discussed here was brought to our attention by David Wiley. The work was supported by Grant No. NIE-G-83-0011 of the Office for Educational Research and Improvement, Center for Education Statistics, and Contract No. N00014-88-K-0304, R&T 4421552 from the Cognitive Science Program, Cognitive and Neural Sciences Division, Office of Naval Research. It does not necessarily reflect the views of either agency. We are grateful to Al Beaton, Martha Stocking and Rebecca Zwick for their comments and suggestions.

Some Consequences of the Uncertainty in

IRT Linking Procedures

Abstract

In many practical applications of item response theory, the parameters of overlapping subsets of test items are estimated from different samples of examinees. A linking procedure is then employed to place the resulting item parameter estimates onto a common scale. It is standard practice to ignore the uncertainty associated with the linking step when drawing inferences that involve items from different subsets, a situation that arises, for example, in the measurement of change. This paper outlines how the uncertainty can be accounted for, and exemplifies the ideas with a jackknife approximation for the Stocking-Lord linking procedure. Examples from the National Assessment of Educational Progress suggest that the resulting uncertainty will usually be negligible for inferences about individuals, but can constitute a major source of estimation error in aggregate statistics such as changes in group means.

Keywords: Item Response Theory
Linking Transformations
The Stocking-Lord Linking Procedure

1.0 Introduction

A widely cited advantage of item response theory (IRT) in educational measurement is its capability to provide proficiency estimates on a common scale when different examinees are administered different items, or when examinees are administered different items at different points in time. A common practice is to estimate the parameters of a large number of test items, treat the estimates as known true parameters, and calculate proficiency estimates for individuals or groups based on responses to selected subsets of items. Practical considerations often preclude administering all items to a single sample of examinees in order to obtain the initial item parameter estimates; rather, estimates for overlapping sets of items are obtained from separate samples of examinees, then linked to a common scale. While it is generally recognized that the parameters of the required linking functions used in practice are estimates rather than known constants, the effects of the uncertainty associated with them upon subsequent analyses are rarely taken into account.

This paper lays out a framework for incorporating the uncertainty associated with IRT linking procedures in subsequent estimates of individual or group change. The ideas are implemented for the linking procedure given by Stocking and Lord (1983), and illustrated with data from the 1984 and 1986 reading surveys of the National Assessment of Educational Progress.

2.0 The 3-Parameter Logistic Item Response Model

The 3PL model expresses the probability of a correct response to an item as a function of (i) the examinee's proficiency level θ_i, and (ii) three parameters characterizing the item, $\beta_j=(a_j,b_j,c_j)$ for $j=1,...,n$. The parameter a_j, called the discrimination or slope parameter, characterizes the item's sensitivity to proficiency. The parameter b_j, called the threshold parameter, is a measure of item difficulty. The parameter c_j is the probability that an individual with very low proficiency will respond correctly to the item. The conditional probability of a correct response to any single item, denoted $P_j(\theta_i)$, is obtained as

$$P(x_{ij}=1|\theta_i,\beta_j) = P(x_{ij}=1|\theta_i,a_j,b_j,c_j)$$

$$= c_j + (1-c_j)/\{1+\exp[-1.7a_j(\theta_i-b_j)]\} \quad , \qquad (1)$$

where the item response $x_{ij} = 1$ if correct and 0 if not. Under the usual assumption of local or conditional independence, the probability of a vector of observed item responses, $\mathbf{x}_i = (x_{i1},...x_{in})$, given a known proficiency value θ_i, can be expressed as a product over items as follows

$$P(\mathbf{x}_i|\theta_i,\beta) = \prod_{j=1}^{n} P(x_{ij}=1|\theta_i,\beta_j)^{x_{ij}}(1-P(x_{ij}=1|\theta_i,\beta_j))^{1-x_{ij}}$$

$$= \prod_{j=1}^{n} P_j(\theta_i)^{x_{ij}} (1-P_j(\theta_i))^{1-x_{ij}} \quad . \qquad (2)$$

Because $P_j(\theta_i)$ is defined as a function of $a_j(\theta_i - b_j)$, the origin and unit of measurement of the proficiency metric are undetermined. That is, for any rescaling constants A and B, if $\theta_i^* = A\,\theta_i + B$, $b_j^* = A\,b_j + B$ and $a_j^* = A^{-1}\,a_j$, then $a_j^*(\theta_i^* - b_j^*) = a_j(\theta_i - b_j)$ and $P_j(\theta_i)$ is unchanged. Since any such linear transformation of the scale retains the meaning and the implications of all parameter values, the unit-size and origin of the θ scale must be determined arbitrarily by the researcher.

Two widely used procedures for estimating the item parameters $\beta=(\beta_1,\ldots,\beta_n)$ of n items under the 3PL model are: joint maximum likelihood, the approach incorporated in the LOGIST program (Wingersky, Barton, and Lord, 1982); and marginal maximum likelihood, the approach incorporated in the BILOG program (Mislevy and Bock, 1982). In both of these programs, the aforementioned linear indeterminacy is resolved by standardizing the distribution of proficiency in the calibration sample in one way or another. The resulting item parameter estimates, and the scale they implicitly define, are then typically taken as fixed when used to estimate individual examinees' proficiencies (as may be required for selection or placement decisions) or population characteristics such as group means (as may be required in educational surveys such as NAEP). In order to focus attention on the impact of the uncertainty in the linking functions, we shall not deal with the uncertainty in the item parameter estimates themselves. The interested reader is referred to Lewis (1985) and Tsutakawa (1986) for more on this latter topic.

3.0 Linking Transformations

Often, it is not feasible to administer all of the items in a large item pool to a single sample of examinees. Instead, overlapping subsets of items are administered to different samples of examinees. When practical considerations preclude a concurrent calibration of all sample data together, as may be the case when the various samples are collected at different points in time, then independent calibrations must be performed on the data collected from each sample. If the IRT model is true, the parameter estimates obtained for items common to two or more calibrations will differ by (i) estimation error, and (ii) an unknown linear transformation.

In this paper, we address the simple case of two tests that share a subset of common items. Each test is independently calibrated on a different sample of examinees. The two calibration samples could represent the same group of examinees tested at two different points in time, or two different groups of examinees for which comparisons are to be made. We refer to the scale established by the calibration of the first sample as the *target* scale and the scale established by the calibration of the second sample as the *provisional* scale. The inferential problems are, first, to estimate the linear transformation needed to bring the item parameter and proficiency estimates from the provisional scale to the target scale, and second, to account for the uncertainty of the linking procedure when stating the precision of resulting statistics. This simple case can be generalized to the more complex calibration problem which arises when multiple forms of a test are calibrated on several independent samples of examinees.

3.1 The Stocking-Lord Linking Procedure

A number of approaches have been suggested for estimating linking transformations. Several attempt to match characteristics of the distributions of a and b parameter estimates on the target scale and reexpressed scale (e.g., Marco, 1977), possibly with differential weighting of estimates to account for the precision with which they have been estimated (Linn, Levine, Hastings, and Wordrop, 1980) or to discount the influence of outliers (Bejar and Wingersky, 1981). The Stocking-Lord (1983) procedure, which we employ in the sequel, minimizes the average squared difference between test characteristic curves (TCCs) estimated from the two sets of item parameters available for the common items.

The input data to the Stocking-Lord procedure consists of two sets of parameter estimates for the common items, one set expressed on the target scale and one set expressed on the provisional scale. For item j, we denote these estimated parameters as $(\hat{a}_{j1},\hat{b}_{j1},\hat{c}_{j1})$ and $(\hat{a}_{j2p},\hat{b}_{j2p},\hat{c}_{j2p})$ respectively.

The goal is to estimate the parameters A and B of the linking transformation that can be used to produce rescaled parameter estimates $(\hat{a}_{j2r}, \hat{b}_{j2r}, \hat{c}_{j2r})$, where

$$\hat{a}_{j2r} = A^{-1} \hat{a}_{j2p} ,$$
$$\hat{b}_{j2r} = A \hat{b}_{j2p} + B , \text{ and}$$
$$\hat{c}_{j2r} = \hat{c}_{j2p} .$$

(Note that the estimate of the lower asymptote parameter $\hat{c}_{j2p}$ is unaffected by the transformation.) After A and B have been estimated from the items common to both calibrations, this same linking transformation is applied to the parameters of the items that appeared in the second calibration only, in order to bring them to the target scale.

Estimation of A and B is accomplished by minimizing the squared difference between estimated true scores (expected numbers correct) on the n_c common items at N preselected values of θ. The function to be minimized is

$$f(A, B, \theta) = 1/N \sum_{i=1}^{N} (\zeta_1(1, 0, \theta_i) - \zeta_2(A, B, \theta_i))^2 \qquad (3)$$

where $\zeta_1(1, 0, \theta_i)$ is the true score associated with the proficiency level θ_i, calculated from the common items using the item parameter estimates expressed on the target scale, and $\zeta_2(A, B, \theta_i)$ is the true score associated with the proficiency level θ_i, calculated from the common items using the item parameter estimates which were originally obtained on the provisional scale and then reexpressed on the target scale with the rescaling parameters A and B. That is,

$$\zeta_1(1,0,\theta_i) = \sum_{j=1}^{n_c} \hat{c}_{j1} + (1-\hat{c}_{j1})/\{1+\exp[-1.7\hat{a}_{j1}(\theta_i-\hat{b}_{j1})]\}$$

and

$$\zeta_2(A,B,\theta_i) = \sum_{j=1}^{n_c} \hat{c}_{j2p} + (1-\hat{c}_{j2p})/\{1+\exp[-1.7A^{-1}\hat{a}_{j2p}(\theta_i-(A\hat{b}_{j2p}+B))]\}$$

$$= \sum_{j=1}^{n_c} \hat{c}_{j2r} + (1-\hat{c}_{j2r})/\{1+\exp[-1.7\hat{a}_{j2r}(\theta_i-\hat{b}_{j2r})]\} .$$

The values $\theta=(\theta_1,...,\theta_N)$, which are selected rather than estimated, play the role of the independent variables in a regression analysis. They should be selected to insure that the equation given in (3) is minimized over the entire (expected) range of the target proficiency scale.

We note in passing that under this procedure, the common items end up with three sets of item parameter estimates, one set expressed on the provisional scale, and two sets expressed on the target scale. Alternative procedures for combining the two sets of estimates expressed on the target scale are given in McKinley (1988).

3.2 A Jackknife Approximation for the Uncertainty of the Stocking-Lord Linking Procedure

The uncertainty associated with the estimated rescaling parameters A and B of the Stocking-Lord linking procedure can be approximated using a Jackknife procedure (Mosteller and Tukey, 1977). Although alternative Jackknife implementations may be appropriate for the problem described here, for the purposes of illustration, we present a single variation only. The variation presented is an example of an interpenetrating Jackknife procedure. It consists of three steps. First, the set of n_c common items used to define the transformation are divided into ten equal length subsets with approximately equal average difficulty. Second, the function given in (3) is minimized ten times. Each minimization is accomplished using all but one of the item subsets defined in step 1. Finally, the observed variation among the A and B parameter estimates obtained from the ten minimizations is used to estimate a covariance matrix which quantifies uncertainty due to (i) the imprecision of the estimated item parameters, and (ii) lack of fit from the IRT model. This procedure is illustrated with data from the National Assessment of Educational Progress in Section 5.

The jackknife procedure described above measures variation arising from two sources: estimation error and model misfit. The uncertainty associated with estimation error can often be decreased by increasing the size of the calibration samples. To decrease the uncertainty associated with model misfit, it is also necessary to have a large number of linking items. To see this, note that, if the IRT model were correct, the differences between sets of (a,b,c) estimates obtained from different increasingly large samples of examinees would be accounted for totally by a linear transformation. In this case, consistent estimates of the linking parameters could be obtained with as few as two linking items. When the IRT model does not fit, however, different sets of linking items will tend to provide different estimates of the linking parameters even as calibration sample sizes increase without bound. In this latter case, it is clear that the model misfit component of uncertainty can only be reduced by increasing the number of linking items. Moreover, the linking items should be chosen so as to be representative of the set of all items which might have been used to estimate the linking function.

4. How the Uncertainty in Linking Procedures Propagates to Subsequent Analyses

In this section, we show how the uncertainty associated with an IRT linking procedure can be accounted for, in the context of measuring change. As before, we consider the simple case of only two tests sharing a single subset of common items The first test is administered to a group of examinees at time 1. The second test is administered to the same group of examinees at time 2. Our primary interest is to measure the change in proficiency observed over time for individual examinees and for specified population subgroups. We assume that a covariance matrix quantifying the uncertainty associated with the parameters of the linear transformation used to link the two tests has been estimated (as with a jackknife approximation, for example).

We first consider the problem of estimating the change in proficiency for a single examinee. Let $\hat{\theta}_{i1}$ denote a proficiency estimate calculated for the ith examinee at time 1 using the estimated item parameters which were originally obtained on the target scale. Let $\hat{\theta}_{i2p}$ denote a proficiency estimate calculated for the same examinee at time 2 using the estimated item parameters which were originally expressed on the provisional scale. And finally, let $\hat{\theta}_{i2r}$ denote a proficiency estimate obtained for the same examinee at time 2 using the item parameters which were originally estimated on the provisional scale and subsequently reexpressed on the target scale; that is, $\hat{\theta}_{i2r} = A\,\hat{\theta}_{i2p} + B$. Since $\hat{\theta}_{i1}$ and $\hat{\theta}_{i2r}$ are both expressed on the target scale, an estimate of the change in proficiency for this examinee can be obtained from the difference, $\hat{D}_i = \hat{\theta}_{i2r} - \hat{\theta}_{i1}$. If the parameters of the linking transformation were known without error, then the standard error of this estimated change would be given by

$$SE(\hat{D}_i) = SE(\hat{\theta}_{i2r} - \hat{\theta}_{i1}) = (\sigma^2_{i2r} + \sigma^2_{i1})^{1/2} , \qquad (4)$$

where σ_{i2r} and σ_{i1} are the standard errors of the proficiency estimates $\hat{\theta}_{i2r}$ and $\hat{\theta}_{i1}$, respectively. (As is usually the case, we have also assumed independent errors across tests.)

Now σ_{i1} will be a function of the item parameters which were originally estimated on the target scale, whereas σ_{i2r} will be a function of the item parameters which were originally estimated on the provisional scale and then reexpressed on the target scale. Thus, any procedure which accounts for the uncertainty of the transformation used to link the two tests will affect the calculation of σ_{i2r} but not σ_{i1}. To calculate σ_{i2r}, note that $\hat{\theta}_{i2r} = \hat{A}\,\hat{\theta}_{i2p} + \hat{B}$, and that the estimated standard error of $\hat{\theta}_{i2p}$, denoted σ_{i2p}, can be calculated as a function of item parameters which have not yet been rescaled and are thus unaffected by the uncertainty of the linking procedure.

As a first step, define a covariance matrix for $[\hat{\theta}_{i2p}, \hat{A}, \hat{B}]$ as follows:

$$\Sigma = \begin{bmatrix} \sigma^2_{i2p} & 0 & 0 \\ 0 & \sigma^2_A & \sigma_{AB} \\ 0 & \sigma_{AB} & \sigma^2_B \end{bmatrix}$$

where σ_A, σ_B, and σ_{AB} quantify estimation variation for the parameters A and B of the linking transformation. The quantities σ_A, σ_B, and σ_{AB} can be approximated using the jackknife procedure given in the previous section. Second, note that

$$\begin{aligned} \mathrm{Var}(\hat{\theta}_{i2r}) &= \mathrm{Var}(\hat{A}\hat{\theta}_{i2p} + \hat{B}) \\ &= \mathrm{Var}(g(\hat{\theta}_{i2p}, \hat{A}, \hat{B})) \\ &\approx \begin{bmatrix} \frac{\partial(g)}{\partial\theta_{i2p}}, & \frac{\partial(g)}{\partial A}, & \frac{\partial(g)}{\partial B} \end{bmatrix} \Sigma \begin{bmatrix} \frac{\partial(g)}{\partial\theta_{i2p}}, & \frac{\partial(g)}{\partial A}, & \frac{\partial(g)}{\partial B} \end{bmatrix}' \\ &= [\, A \,,\, \theta_{i2p} \,,\, 1 \,]\, \Sigma\, [\, A \,,\, \theta_{i2p} \,,\, 1 \,]' \end{aligned}$$

$$\approx \hat{A}^2\hat{\sigma}^2_{i2p} + \hat{\theta}^2_{i2p}\hat{\sigma}^2_A + 2\hat{\theta}_{i2p}\hat{\sigma}_{AB} + \hat{\sigma}^2_B$$

$$\approx f(\hat{\theta}_{i2p},\hat{A},\hat{\Sigma}) \quad . \qquad (5)$$

Thus, the uncertainty associated with the linking procedure can be accounted for in the estimated standard error of the difference D_i, as follows:

$$SE(\hat{D}_i) = SE(\hat{\theta}_{i2r} - \hat{\theta}_{i1})$$

$$= \left(Var(\hat{\theta}_{i2r}) + \sigma^2_{i1} \right)^{1/2}$$

$$= \left(f(\hat{\theta}_{i2p},\hat{A},\Sigma) + \sigma^2_{i1} \right)^{1/2} \qquad (6)$$

where $f(\hat{\theta}_{i2p},\hat{A},\Sigma)$ is given as in (5).

The same procedure can also be used to incorporate the uncertainty associated with the linking parameters A and B in the estimated standard error of aggregate statistics such as the difference between two subgroup means. In this latter case, the θ and σ statistics for individuals will be replaced by corresponding point estimates and standard errors for subgroup means.

5. A Numerical Illustration

In this section, data available from the National Assessment of Educational Progress (NAEP), a congressionally mandated survey of the educational achievement of American students, is used to approximate the uncertainty of the Stocking-Lord linking procedure and to evaluate the consequences of that uncertainty. Data from two NAEP surveys are used: the 1984 Reading Survey and the 1986

Reading Survey. Both of these surveys were independently scaled using a three parameter logistic IRT model. Item parameters were estimated using BILOG (Mislevy & Bock, 1982) and mean proficiencies for population subgroups were obtained using the plausible values methodology given in Mislevy and Sheehan (1987). These data are used to illustrate the consequences of the uncertainty of the transformation parameter estimates from the Stocking-Lord linking procedure. Because NAEP data support inferences about aggregate statistics such as group means but not about individuals' proficiencies, we use real NAEP data to demonstrate procedures for changes in group means but simulated data for changes in individual proficiencies.

5.1 The NAEP Data

Mean reading proficiencies for the three age groups which were assessed by NAEP in 1984 and 1986 are given in Table 1. The first row of the table provides 1984 age group means expressed on the 1984 calibration scale. For the purpose of this illustration, the 1984 calibration scale is designated as the target scale. The second and third rows of the table provide 1986 age group means expressed on the provisional scale (the 1986 calibration scale) and the target scale (the 1984 calibration scale). The Stocking-Lord linking procedure was used to estimate the linear transformation needed to express the 1986 means on the 1984 calibration scale. The table also provides estimated standard errors for each mean.

Table 1 about here

5.2 Quantifying the Uncertainty of the NAEP Link

The 1984 NAEP survey contained 128 cognitive reading items. The 1986 NAEP survey contained 107 cognitive reading items, 76 which were common to the 1984 assessment and 31 which were administered for the first time in 1986. The linking transformation needed to express the item parameters obtained from the calibration of the 1986 data on the scale established by the calibration of the 1984 data was estimated using the Stocking-Lord linking procedure, as implemented in the TBLT computer program (Stocking, 1986). The generally satisfactory results can be seen in Figure 1, which shows the TCCs of the first and second calibrations of the common items after reexpression, and in Figure 2, which plots the b-parameter estimates from the first and reexpressed second calibrations. The jackknife procedure described in Section 3 was used to approximate the uncertainty associated with the estimated parameters of the linking transformation. The results are given in Table 2.

Figures 1 and 2 and Table 2 about here

5.3 Inference for a Single Examinee

The artificial data set constructed for this analysis contained simulated responses for five examinees to two tests. The first test consisted of 30 items selected from the 1984 NAEP reading survey. The second test consisted of 30 items selected from the 1986 NAEP reading survey, half of which were common to the 1984 survey. For a given examinee, responses were generated in accordance with the 3PL, with item parameter estimates for the first test taken from the 1984 NAEP calibration run and item parameter estimates for the second test taken from the 1986 NAEP calibration run. So that the proficiency of a given simulee was the same on both tests, a value of θ was specified for the first test and $(\theta-B)/A$ was used for the second. Simulees' θ values on the first test were -1.0, -0.5, 0.0, 0.5, and 1.0. The response vectors generated according to these specifications are given in Table 3.

Table 3 about here

Treating the item parameter estimates as known, maximum likelihood estimates (MLEs) of θ and associated standard errors were obtained for each response pattern using the BILOG program. They are shown in Table 4, with the values for the second test shown before and after reexpression. Table 5 provides estimated standard errors for the change from the first test to the second using (4), which does not take the uncertainty of A and B into account, and (6), which does. The increase in standard errors is negligible, about 2-percent on the average. An approximate variance components analysis is given in Table 6. For each response pattern considered, the total error variance is estimated using (6) which includes components due to both sampling and linking. The contribution due to sampling alone is estimated using (4) and the contribution due to linking is obtained by subtraction. The table shows that for each response pattern considered, the relative increase in uncertainty is negligible, accounting for about three percent of the total error variance on the average.

Tables 4, 5 and 6 about here

5.4 Inference for Group Means

The changes in the mean reading proficiencies of students aged 9, 13 and 17, over the two year period from 1984 to 1986, as estimated from the NAEP data, are given in Table 7.[1] The table also provides approximate standard errors calculated using (4) and (6). Whereas the size of standard errors increased by only about 2-percent for estimates of change of individuals, the increase in standard errors for groups is about 200-percent! An approximate

[1] These figures are shown for illustrative purposes only, and are not to be taken as estimates of changes in reading proficiency during the period due to certain anomalies in the 1985/86 NAEP data. The interested reader is referred to Beaton (1988) for further information.

variance components analysis is given in Table 8. The table shows that the component due to linking represents approximately 90-percent of the total error variance, on the average. To put these results in another perspective, the change in mean reading proficiency at each age level is expressed in standard error units in Table 9. The table shows, for example, that the decrease in the mean reading proficiency of 9 year olds is approximately three standard errors when the uncertainty of the linking procedure is not accounted for, but only one standard error when it is.

Tables 7, 8 and 9 about here

6.0 Summary

A common problem in applied work with item response theory is to express item parameter estimates from separate calibrations on the same scale. based on the multiple estimates for subsets of items common to two or more calibrations. Several methods have been proposed for estimating the optimal linear transformations for this purpose, including the Stocking-Lord (1983) procedure for matching test characteristic curves. After the resulting transformations have been applied, the uncertainty associated with them is rarely taken into account in subsequent analyses of individual or group levels of proficiency.

This uncertainty can be expressed in terms of a covariance matrix of estimation errors, which can be approximated empirically through a procedure such as the jackknife. With an approximation of the sampling covariance matrix of estimation errors of the parameters of a linking transformation, one can readily derive standard errors for change scores or comparisons that take this additional uncertainty into account.

Using data from the 1984 and 1986 reading surveys of the National Assessment of Educational Progress, this paper used the jackknife to approximate the uncertainty of the linking transformation between the two assessments. Its effect was found to be negligible in the context of drawing inferences about change of individuals, since its magnitude was much smaller than the uncertainty arising from having only the limited numbers of item responses from individuals that generally characterize individual testing programs. Correct standard errors were only about 2-percent larger than those that ignored linking uncertainty. The effect was substantial in the context of estimating group changes, however, leading to correct standard errors that were 200-percent larger. The differential impact is due to the fact that sampling variances of group means are much smaller than sampling variances of individual scores, while the sampling variance of the linking transformation is the same in both cases.

References

Beaton, A.E. (1988). **The NAEP 1985-86 Reading Anomaly: A Technical Report.** Princeton, NJ: Educational Testing Service.

Bejar, I., and Wingersky, M.S. (1981). An application of item response theory to equating the Test of Standard Written English. College Board Report 81-1. Princeton, NJ: Educational Testing Service.

Lewis, C.E. (1985). Estimating individual abilities with imperfectly known item response functions. Paper presented at the 50th Anniversity Meeting of the Psychometric Society, Nashville, TN, June 1-4.

Linn, R.L., Levine, M.V., Hastings, C.N., and Wordrop, J.L. (1980). An investigation of item bias in a test of reading comprehension. Technical Report No. 163. Urbana, IL: Center for the Study of Reading, University of Illinois.

Marco, G.L. (1977). Item characteristic curve solutions to three intractible testing problems. **Journal of Educational Measurement**, 14, 1139-160.

McKinley, R.L. (1988). A comparison of six methods for combining multiple IRT Item Parameter Estimates. **Journal of Educational Measurement, 25**, nn-nn.

Mislevy, R.J., and Bock, R.D. (1982). **BILOG: Item analysis and test scoring with binary logistic models** [computer program]. Mooresville, IN: Scientific Software, Inc.

Mislevy, R.J., and Sheehan, K.M. (1987). Marginal estimation procedures. In A.E. Beaton, **Implementing the new design: The NAEP 1983-84 technical report**. (Report No. 15-TR-20). Princeton, NJ: Educational Testing Service.

Mosteller, F., and Tukey, J.W. (1977). **Data Analysis and Regression**. Reading, MS: Addison-Wesley.

Stocking, M.L. (1986) **TBLT: linking tests by matching test response curves** [computer program]. Princeton, NJ: Educational Testing Service.

Stocking, M.L., and Lord, F.M. (1983). Developing a common metric in item response theory. **Applied Psychological Measurement**, **7**, 201-210.

Tsutakawa, R.K. (1986). Approximation for Bayesian ability estimation. Paper presented at the Office of Naval Research contractors' conference on Model-Based Measurement, Gatlinburg, TN, April 27-30.

Wingersky, M.S., Barton, M.A., and Lord, F.M. (1982). **LOGIST user's guide**. Princeton, NJ: Educational Testing Service.

Table 1

Mean Proficiencies
Estimated from the 1984 and 1986 NAEP Reading Surveys
With Standard Errors in Parentheses

Year	Scale	Age 9	Age 13	Age 17
84	84 Calib.	-0.752(.020)	0.150(.014)	0.766(.018)
86	86 Calib.	-0.375(.025)	0.571(.019)	0.874(.018)
86	84 Calib.	-0.864(.028)	0.198(.022)	0.538(.020)

The 1984 sample included over 22,000 students at each age level. The 1986 sample included approximately 7,000 Age 9 students, 6,000 Age 13 students, and 16,000 Age 17 students.

Table 2

Results of the Jackknife Approximation for the Stocking-Lord Linking Procedure

Run[1]	Items	A	B
0	76	1.122196	-0.442910
1	68	1.118018	-0.449670
2	68	1.126296	-0.447837
3	68	1.121856	-0.449472
4	68	1.110982	-0.433893
5	68	1.114703	-0.426793
6	68	1.128065	-0.430320
7	69	1.125834	-0.446748
8	69	1.128753	-0.440663
9	69	1.112862	-0.447648
10	69	1.135424	-0.455858

Parameter	Jackknife Estimate
σ^2_A	0.00512
σ^2_B	0.00740
σ_{AB}	-0.00238

[1] The parameter estimates, A and B, obtained from Run 0 were used to reexpress the 1986 results on the 1984 scale. The parameter estimates obtained from Runs 1 through 10 were used only to estimate the uncertainty of the linking procedure.

Table 3

Simulated Responses To Test 1
Administered at Time 1

Generating Value	
-1.0	11000 11000 10011 00101 00000 01010
-0.5	00110 10101 10000 10011 00111 11001
0.0	00010 11101 11100 00100 01110 11100
0.5	11111 01111 11111 00111 01101 11111
1.0	11111 11111 11111 01111 10110 11111

Simulated Responses To Test 2
Administered at Time 2

Generating Value	
-.50	00010 01000 00011 11000 10000 00001
-.05	11001 01000 01011 11101 01100 11000
0.39	01100 01101 10011 00111 11111 10100
0.84	00011 11111 10111 11110 11101 01111
1.29	11111 11111 11111 10111 10110 01110

Table 4

Maximum Likelihood Estimates of Reading Proficiency
At Time 1 and Time 2
For Five Simulated Subjects
With Estimated Standard Errors in Parentheses

Generating Value	Value Estimated at Time 1	Value Estimated at Time 2: Before Reexpression	Value Estimated at Time 2: After Reexpression
-1.0	-1.062 (.625)	-0.375 (.422)	-0.864 (.474)
-0.5	-0.662 (.489)	-0.116 (.534)	-0.574 (.560)
0.0	-0.502 (.470)	0.249 (.360)	-0.163 (.404)
0.5	0.748 (.546)	0.824 (.409)	0.482 (.459)
1.0	1.177 (.662)	1.434 (.512)	1.172 (.574)

Table 5

An Estimate of the Change in Reading Proficiency
From Time 1 to Time 2
For Five Simulated Subjects
With Approximate Standard Errors[1]

Change in Generating Values	Estimated Change	S.E. Method 1	S.E. Method 2
0	0.198	0.784	0.790
0	0.088	0.743	0.779
0	0.339	0.620	0.625
0	-0.266	0.713	0.718
0	-0.005	0.876	0.883

[1] Method 1 refers to the method which assumes that the linking function is known without error, as in equation (4); Method 2 refers to the method which accounts for the uncertainty of the linking procedure as in equation (6).

Table 6

A Comparison of Approximate Variance Components
For Inferences About Change at the Individual Level

Generating Value	Total Variance	Component Due to Sampling	Component Due to Linking	Linking Variance as % of Total Variance
-1.0	.6241	.6146	.0094	1.5
-0.5	.6068	.5520	.0548	9.0
0.0	.3906	.3844	.0062	1.6
0.5	.5155	.5084	.0071	1.4
1.0	.7797	.7674	.0123	1.6

Table 7

An Estimate of the Change in Mean Reading Proficiency
From 1984 to 1986
With Approximate Standard Errors[1]

Age	Estimated Change	S.E. Method 1	S.E. Method 2
9	-0.112	.034	.105
13	0.048	.026	.084
17	-0.228	.027	.066

[1] Method 1 refers to the method which assumes that the linking function is known without error, as in equation (4); Method 2 refers to the method which accounts for the uncertainty of the linking procedure as in equation (6).

Table 8

A Comparison of Approximate Variance Components
For Inferences About Change at the Group Level

Age	Total Variance[1]	Component due to Sampling	Component due to Linking	Linking Variance as % of Total Variance
9	.0110	.0012	.0098	89.5
13	.0071	.0007	.0064	90.1
17	.0044	.0007	.0037	84.1

1 Total Variance refers to the estimated variance of the change in mean reading proficiency from 1984 to 1986.

Table 9

The Estimated Change in Mean Reading Proficiency
from 1984 to 1986
Expressed in Standard Error Units

Age	Method 1 S.E. Units	Method 2 S.E. Units
9	-3.29	-1.07
13	1.85	0.57
17	-8.44	-3.45

Figure 1

Comparison of Test Characteristic Curves

Solid Line = 1984 Curve

Dashed Line = Reexpressed 1986 Curve

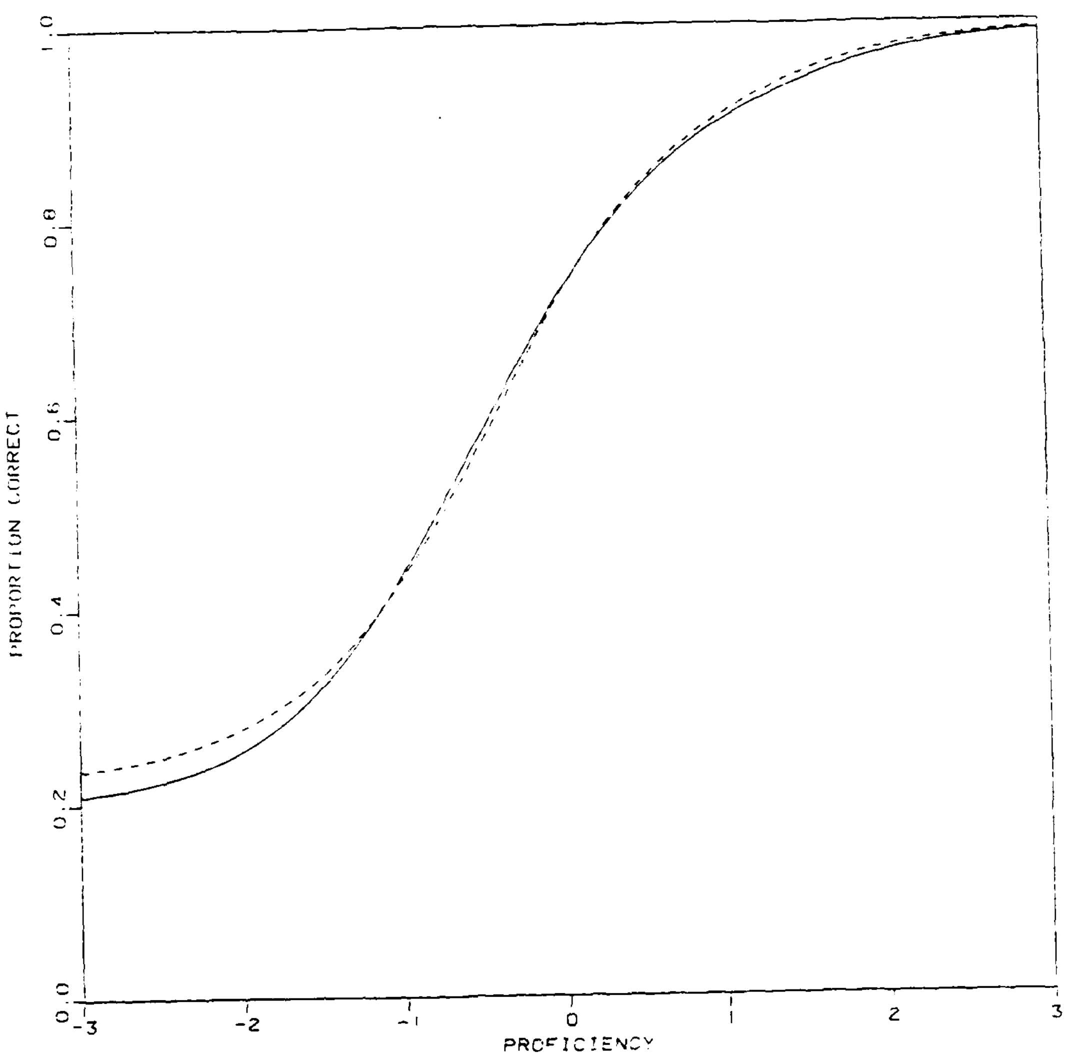
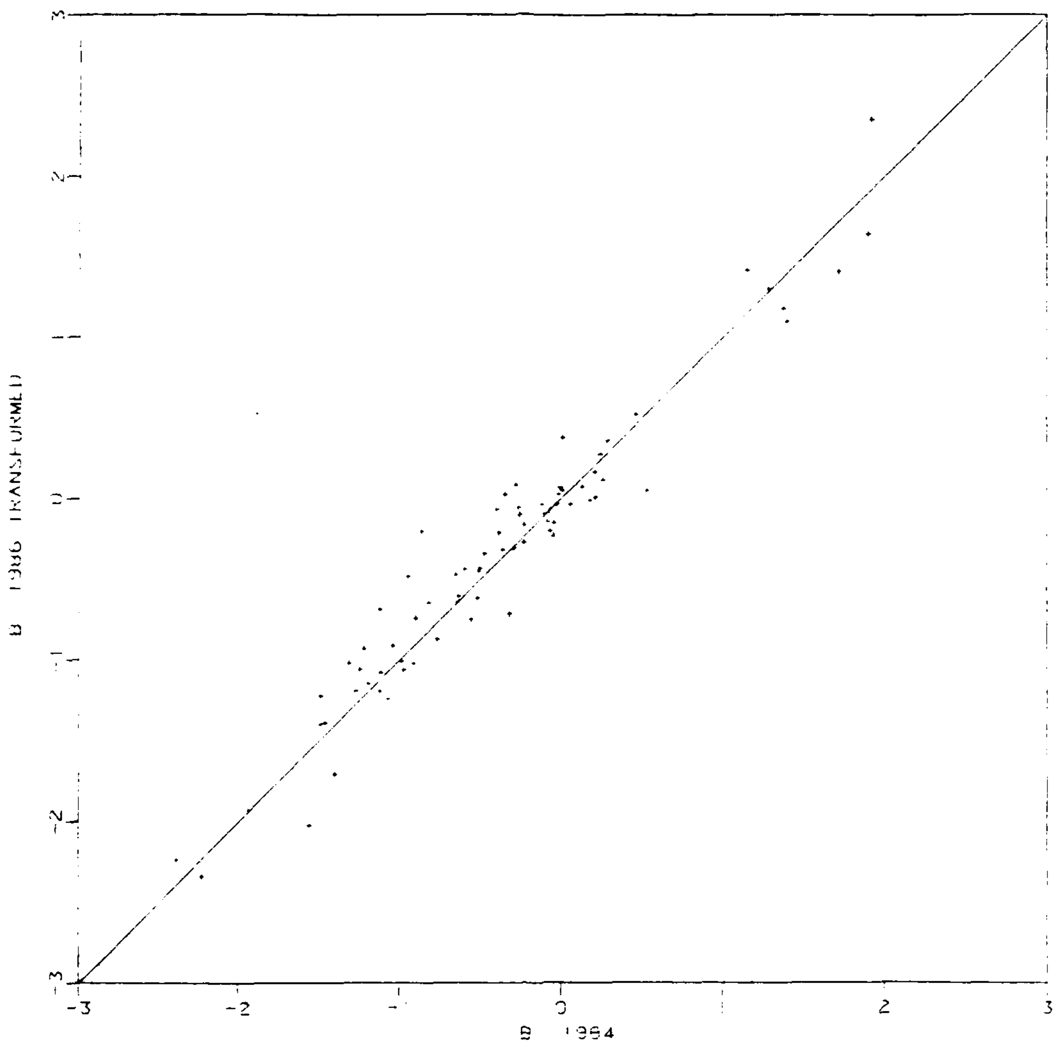


Figure 2

Comparison of Item b Parameter Estimates

Reexpressed 1986 Estimates vs. 1984 Estimates



Dr. Terry Ackerman
American College Testing Programs
P.O. Box 168
Iowa City, IA 52243

Dr. Robert Ahlers
Code N711
Human Factors Laboratory
Naval Training Systems Center
Orlando, FL 32813

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

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

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

Dr. Isaac Bejar
Mail Stop: 10-R
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

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

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

Dr. Bruce Bloxom
Defense Manpower Data Center
550 Camino El Estero,
Suite 200
Monterey, CA 93943-3231

Dr. R. Darrell Bock
University of Chicago
NORC
6030 South Ellis
Chicago, IL 60637

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

Dr. Robert Breaux
Code 7B
Naval Training Systems Center
Orlando, FL 32813-7100

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

Dr. James Carlson
American College Testing
Program
P.O. Box 168
Iowa City, IA 52243

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

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

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

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

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

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

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

Dr. Timothy Davey
Educational Testing Service
Princeton, NJ 08541

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

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

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

Dr. Hei-Ki Dong
Bell Communications Research
6 Corporate Place
PYA-1K226
Piscataway, NJ 08854

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

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

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

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

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

Dr. John M. Eddins
University of Illinois
252 Engineering Research
Laboratory
103 South Mathews Street
Urbana, IL 61801

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

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

Dr. Benjamin A. Fairbank
Performance Metrics, Inc.
5825 Callaghan
Suite 225
San Antonio, TX 78228

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Prof. Lutz F. Hornke
Institut fur Psychologie
RWTH Aachen
Jaegerstrasse 17/19
D-5100 Aachen
WEST GERMANY

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

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

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

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

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

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

Dr. Douglas H. Jones
Thatcher Jones Associates
P.O. Box 6640
10 Trafalgar Court
Lawrenceville, NJ 08648

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

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

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

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

Dr. James Kraatz
Computer-based Education
Research Laboratory
University of Illinois
Urbana, IL 61801

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

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

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

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

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

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

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

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

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

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

Dr. James R. McBride
The Psychological Corporation
1250 Sixth Avenue
San Diego, CA 92101

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

Dr. Robert McKinley
Educational Testing Service
16-T
Princeton, NJ 08541

Dr. James McMichael
Technical Director
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Barbara Means
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

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

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

Headquarters Marine Corps
Code MPI-20
Washington, DC 20380

Dr. W. Alan Nicewander
University of Oklahoma
Department of Psychology
Norman, OK 73071

Deputy Technical Director
NPRDC Code 01A
San Diego, CA 92152-6800

Director, Training Laboratory,
NPRDC (Code 05)
San Diego, CA 92152-6800

Director, Manpower and Personnel
Laboratory,
NPRDC (Code 06)
San Diego, CA 92152-6800

Director, Human Factors
& Organizational Systems Lab,
NPRDC (Code 07)
San Diego, CA 92152-6800

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

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

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

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

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

Office of Naval Research,
Code 125
800 N. Quincy Street
Arlington, VA 22217-5000

Assistant for MPT Research,
Development and Studies
OP 0187
Washington, DC 20370

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

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

Dr. Randolph Park
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

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

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

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

Department of Operations Research,
Naval Postgraduate School
Monterey, CA 93940

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

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

Dr. Barry Riegelhaupt
HumRRO
1100 South Washington Street
Alexandria, VA 22314

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

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

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

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

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

Dr. Mary Schratz
Navy Personnel R&D Center
San Diego, CA 92152-6800

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

Dr. W. Steve Sellman
OASD(MRA&L)
2B269 The Pentagon
Washington, DC 20301

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

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

Dr. H. Wallace Sinaiko
Manpower Research
and Advisory Services
Smithsonian Institution
801 North Pitt Street, Suite 120
Alexandria, VA 22314-1713

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

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

Dr. Paul Speckman
University of Missouri
Department of Statistics
Columbia, MO 65201

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

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

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

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

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

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

Dr. Kikumi Tatsuoka
CERL
252 Engineering Research
Laboratory
103 S. Mathews Avenue
Urbana, IL 61801

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

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

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

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

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

Educational Testing Service/Mislevy

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

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

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

Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

Dr. Ming-Mei Wang
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. Thomas A. Warm
Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Dr. Brian Waters
HumRRO
12908 Argyle Circle
Alexandria, VA 22314

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

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

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

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

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

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

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

Dr. Hilda Wing
NRC MH-176
2101 Constitution Ave.
Washington, DC 20418

Dr. Martin F. Wiskoff
Defense Manpower Data Center
550 Camino El Estero
Suite 200
Monterey, CA 93943-3231

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

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

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

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

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

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

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

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