The Effect of Item Parameter Estimation Error on Decisions Made Using the Sequential Probability Ratio Test

Research Report ONR 87-1

Judith A. Spray
Mark D. Reckase


Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

September 1987

Best Available Copy
The effect of item parameter estimation error on decisions made using the sequential probability ratio test

A series of computer simulations were performed in order to observe the effects of item response theory (IRT) item parameter estimation error on decisions made using an IRT-based sequential probability ratio test. Specifically, the effects of such error on misclassification rates and the average number of items required for either a mastery (pass) or nonmastery (fail) decision were observed under varied SPRT conditions. These conditions include the a priori or nominal type I (α) and type II (β) error rates, the simple hypotheses tested by the SPRT procedure, and the composition of the item pool (specifically the a, b, and c parameters which characterized the items according to a three-parameter logistic model) used to administer the SPRT. The results of these simulations showed that these SPRT decisions are not greatly affected by this particular level of error in parameter estimates modeled in this study. Misclassification error rates were slightly greater when estimation error in the item parameters was present, but such differences appear to be negligible.
THE EFFECT OF ITEM PARAMETER ESTIMATION ERROR ON DECISIONS MADE USING THE SEQUENTIAL PROBABILITY RATIO TEST

Judith A. Spray
Mark D. Reckase

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.
ABSTRACT

A series of computer simulations were performed in order to observe the effects of item response theory (IRT) item parameter estimation error on decisions made using an IRT-based sequential probability ratio test. Specifically, the effects of such error on misclassification rates and the average number of items required for either a mastery (pass) or nonmastery (fail) decision were observed under varied SPRT conditions. These conditions included the a priori or nominal type I ($\alpha$) and type II ($\beta$) error rates, the simple hypotheses tested by the SPRT procedure, and the composition of the item pool (specifically the $a$, $b$ and $c$ parameters which characterized the items according to a three-parameter logistic IRT model) used to administer the SPRT. The results of these simulations showed that these SPRT decisions are not greatly affected by this particular level of error in parameter estimates modeled in this study. Misclassification error rates were slightly lower and average numbers of items required for a decision were slightly greater when estimation error in the item parameters was present, but such differences appear to be negligible.
The Effect of Item Parameter Estimation Error on Decisions Made Using the Sequential Probability Ratio Test

Wald's (1947) sequential probability ratio testing (SPRT) procedure has been proposed as a technique for making pass-fail or mastery-nonmastery decisions in adaptive testing situations (Reckase, 1983). The SPRT was originally proposed by Wald in order to decide between two simple hypotheses, $H_0$ and $H_1$, or

$$H_0: \theta = \theta_0$$

vs.

$$H_1: \theta = \theta_1,$$

where $\theta$ is an unknown parameter of the distribution of some random variable, $X$. In a cognitive testing situation, the random variable, $X$, is the response to a test item and is usually assumed to be a dichotomous response, correct or incorrect.

In the case of cognitive testing, the random variable, $X$, is assumed to follow a binomial distribution. If $P(\theta_i)$ is the probability that examinee $i$ will respond correctly to any item and $Q(\theta_i) = 1 - P(\theta_i)$ is the probability of an incorrect response from examinee $i$, then (for any single item) the random variable, $X$, represents a single Bernoulli trial and is distributed as $\text{bin}(P(\theta_i), 1)$. Then, let

$$\pi(\theta_i) = \text{Prob}(X = x | \theta = \theta_i) = P(\theta_i)^x Q(\theta_i)^{1-x}$$

where

$$x = \begin{cases} 1, & \text{correct response} \\ 0, & \text{incorrect response} \end{cases}$$
For any single item, the probability of observing $X = x$ under the alternative hypothesis is $\pi(\theta_1)$. Under the null hypothesis, this probability is $\pi(\theta_0)$. The functions, $\pi(\theta_1)$ and $\pi(\theta_0)$ are called likelihood functions of $x$. A ratio of these two functions, $L(x) = \pi(\theta_1)/\pi(\theta_0)$, is called a likelihood ratio.

Two error probabilities, $\alpha$ and $\beta$, can be defined, where

\[
\text{Prob (choosing } H_1 \text{ if } H_0 \text{ is true}) = \alpha
\]

and

\[
\text{Prob (choosing } H_0 \text{ if } H_1 \text{ is true}) = \beta.
\]

Wald (1947) defined two likelihood ratio boundaries using inequalities which involved these error probabilities. These boundaries are $A$ and $B$ where

lower boundary $= B \geq \beta/(1-\alpha)$

and

upper boundary $= A \leq (1-\beta)/\alpha$.

According to Wald's SPRT, trials or items would be observed in sequence, $x_1, x_2, \ldots, x_n$, and following each observation, the likelihood ratio, $L(x_1, x_2, \ldots, x_n)$, would be computed, where

\[
L(x_1, x_2, \ldots, x_n) = \frac{\pi_1(\theta_1) \cdot \pi_2(\theta_1) \cdots \pi_n(\theta_1)}{\pi_1(\theta_0) \cdot \pi_2(\theta_0) \cdots \pi_n(\theta_0)}.
\]

The likelihood function then would be compared to the boundaries, $A$ and $B$. If
then $H_1$ is accepted. If

$$L(x_1, x_2, \ldots, x_n) \geq A,$$

then $H_0$ is accepted. If

$$L(x_1, x_2, \ldots, x_n) \leq B,$$

then another trial is observed, or in the case of cognitive testing, another item is administered.

Once $a$, $B$, and the hypotheses are set prior to testing, the stopping rules of the test (i.e., the boundaries) are defined. Although $a$ and $B$ are determined prior to observing $x$, where $x = (x_1, x_2, \ldots, x_n)$, Wald (1947) pointed out that the actual error rates observed in practice, $a^*$ and $B^*$, would be bounded from above by

$$a^* \leq a/(1-B)$$

and

$$B^* \leq B/(1-a)$$

(see Wald, 1947, p. 46). This means that even though the nominal error probabilities, $a$ and $B$, are established prior to testing, the actual error rates can be less than these nominal rates, or even greater than the nominal rates.
Reckase (1983) reported the results of computer simulation research of the SPRT procedure as it applied to tailored or computerized adaptive testing (CAT) for making mastery testing decisions. He noted that this research had three purposes: (1) to obtain information on how the SPRT procedure functioned when items were selected from the item pools on the basis of maximizing item information rather than on the basis of a simple random sampling procedure; (2) to gain experience in selecting values of \( \theta_0 \) and \( \theta_1 \), assumed to be the two critical values of ability required to be classified as nonmaster or master, respectively; and (3) to obtain information on the effects of guessing on the accuracy of classification when the form of \( P(\theta) \) was the one-parameter logistic IRT (item response theory) model but a three-parameter logistic model was used to determine the responses.

Reckase's first concern, (1) above, was that, in a given pool of test items, only a small portion of these items would be available for selection for a given examinee and that the selection of test items would be based on estimates of \( \theta \) after the administration of, say \( n \) items. This is because the selection of the \( n+1 \)st item is dependent upon maximum item information at \( \hat{\theta}_n \), \( \max I(\hat{\theta}_n) \), where

\[
I(\hat{\theta}_n) = \frac{P'(\hat{\theta}_n)}{P(\theta)Q(\theta_n)},
\]

and \( P'(\hat{\theta}_n) \) is the derivative of \( P(\theta) \) w.r.t. \( \theta \), evaluated at \( \hat{\theta}_n \).

It would appear that this nonrandom selection process would not really be a problem because the stopping rule of the SPRT is determined by prior knowledge of \( \alpha, \beta, \theta_0 \), and \( \theta_1 \) before the test even begins and because \( L(x, x_2, \ldots, x_n) \) is written as the product of the individual item likelihood ratios through the assumption of local independence of the \( x \) given \( \theta_1 \).

However, a problem may occur when it is time to generalize the results of
the mastery/nonmastery decision-making process, as defined by the SPRT. In most mastery situations, it is desirable to generalize the results of a mastery test to the entire domain of objectives measured by the test, and this domain is usually represented by the entire item pool. If, however, items are selected on the basis of \(\max I(\hat{\theta})\), then inferences made to the entire pool of items may be questionable. On the other hand, one could always claim that the inferences are actually being made or generalized to the ability level or the latent trait value (call it \(\theta_c\)) required before an individual examinee can pass the criterion number of items in the item pool, \(n(\theta_c)\).

Perhaps a more serious concern is the effect of assuming that the function, \(P(\theta_i)\), is only a function of \(\theta_i\), and known item parameters. For the IRT models which would be assumed to define \(P(\theta_i)\) explicitly, the item parameters are usually treated as known values in CAT administrations. The item pool contains values of these item parameters so that \(L(x_1, x_2, \ldots, x_n)\) and \(I(\hat{\theta}_n)\) can be computed during the test. However, these values are, themselves, estimates of the true but unknown item parameters. The estimates have been obtained in calibration computer runs prior to the CAT administrations and are stored along with the actual items in the pool.

The present computer simulation study was designed to investigate the effects of item parameter estimation error on the characteristics of the SPRT procedure. In this first phase of a thorough investigation, a strict SPRT was administered, meaning that the test was not adaptive (i.e., \(\theta\) was not estimated and items were not selected for administration based on \(\max I(\hat{\theta})\)).

The research question to be answered by these simulations was, "What are the effects on observed type I (\(\alpha^*\)) and type II (\(\beta^*\)) error rates when an SPRT is administered from item pools which contain items whose parameters are estimates rather than known values?" A secondary interest was to observe the
effects of these conditions on the average number of test items required to
make a classification decision at each value of \( \theta \) (particularly at \( \theta_0 \)
and \( \theta_1 \)). This number, called the average sample number (ASN) is a function of
the stopping rule of the tests (i.e., it is a function of \( a, \beta, \theta_0 \) and \( \theta_1 \)).

Method

Two hundred eighty-eight computer simulations were completed on either an
IBM PC or XT. These 288 simulations represented one combination of conditions
from a 2 \( \times \) 4 \( \times \) 3 \( \times \) 3 \( \times \) 4 completely crossed design. Each of these runs consisted
of 1000 replications of an SPRT administered to all of 24 hypothetical examinees
with ability, \( \theta_i \), ranging from -3.0 to +3.0, incremented by .25.

The research design conditions were (1) an estimation error condition,
(2) composition of the item pools, (3) a priori type I error rate (\( a \)), (4) a
priori type II error rate (\( \beta \)), and (5) hypotheses. It was assumed that the item
pools contained items which interacted with each examinee according to a three-
parameter logistic model (3-PLM) to produce a correct or incorrect response to
each item.

Conditions

Estimation error. There were two levels of the estimation error condi-
tion, absent (E1) or present (E2). Under the absent level (E1), the item
parameters from the items in the pools were considered to be known values, and
each of the 24 hypothetical examinees in the simulations with ability, \( \theta_i \),
responded to the items in the pool by comparing a deviate from a uniform
distribution on the open interval, 0 to 1, with the \( P(\theta_i) \) function given by
the 3-PLM, abbreviated as \( P_i \).
Under the present level, it was assumed that the item parameters were actually estimates derived from previous maximum likelihood estimation (MLE) calibrations on 2500 examinees with ability, $\theta$, distributed as normal with mean zero and variance one. According to the notation used by Thissen and Wainer (1982), the maximum likelihood estimates of the set of item parameters, $\xi$, are those that are located where the partial derivatives of the log of the likelihood function, summed over $N$ examinees, are zero. If $\xi$ is this sum, or

$$
\xi = \sum_{i=1}^{N} x \log (P_i) + (1 - x) \log (1 - P_i),
$$

then, again from Thissen and Wainer (1982) but written without the $i$ subscript, these MLEs satisfy

$$
\frac{\partial \xi}{\partial \xi} = \sum P \frac{\partial P}{\partial \xi} - (1 - x) \frac{\partial P}{\partial \xi} = 0.
$$

(1)

The inverse of the negative expected value of the matrix of second derivatives of the function, $\xi$, is the asymptotic variance-covariance matrix of the estimates, $\hat{\xi}$, obtained from the relationship given by (1). If the second partial derivatives of $\xi$ are written, in general, as $\frac{\partial^2 \xi}{\partial \xi_s \partial \xi_t}$, for any parameters, $\xi_s$ and $\xi_t$, then

$$
-E\left[\frac{\partial^2 \xi}{\partial \xi_s \partial \xi_t}\right] = N \int_{-\infty}^{\infty} \left[ \frac{1}{P} \frac{\partial P}{\partial \xi_s} \frac{\partial P}{\partial \xi_t} + \frac{1}{(1 - P)} \frac{\partial P}{\partial \xi_s} \frac{\partial P}{\partial \xi_t} \right] \phi(\theta) \, d\theta,
$$

(2)

where $\phi(\theta)$ is taken to be a normal density with zero mean and variance one (Thissen & Wainer, 1982). In other words, if $\bar{\xi}$ is the variance-covariance matrix of $\xi$, then $\bar{\xi}$ is defined by the inverse of the matrix whose elements are given by (2).
For the present level (E2) of the estimation error condition, it was assumed that the item parameters were actually estimates sampled from a multivariate normal distribution with mean vector $\xi$ and variance-covariance matrix $\Sigma$, where $\xi$ was given for the item pool used for a particular SPRT and $\Sigma$ was computed from (2).

**Item Pools.** There were four types of item pools used in the simulations. The first three consisted of 500 identical items from a three-parameter logistic IRT model of the form,

$$P(\theta_i) = c + \frac{(1 - c)}{1 + \exp(-1.7a(\theta_i - b))}.$$  \hspace{1cm} (3)

For the first pool (I1), $a = 1$, $b = 0$, and $c = 0$ for all 500 items. Under the $E_1$ condition, these identical items represented a simple SPRT with constant success probability, $P(\theta_i)$ for a given $\theta_i$ value. Under the $E_2$ condition, the items were still administered in sequence but were no longer identical because each item represented a different set of item parameter estimates. For example, even though $a_1 = a_2 = \ldots = a_{500}$, each $a$ parameter represented an estimate, $\hat{a}_j$, where

$$\hat{a}_j = a + \varepsilon_{aj},$$

and $\varepsilon_{aj}$ was a random deviate from a multivariate normal distribution with mean vector $\xi$ and variance-covariance matrix $\Sigma$, defined previously.

For the second item pool (I2), $a = 1$, $b = 0$, and $c = .2$. For the third pool (I3), $a = 1.5$, $b = 0$, and $c = .2$. Again, under $E_1$ these item parameters remained constant for all 500 items in a pool. However, under $E_2$, item parameter values were assumed to be estimates ($a + \varepsilon_{aj}$, $b + \varepsilon_{bj}$, and $c + \varepsilon_{cj}$ with $\varepsilon_{aj}$, $\varepsilon_{bj}$, and $\varepsilon_{cj}$ being random deviates as before).
For the fourth item pool (14), the 500 sets of parameters were generated from a pseudo-random number generator with \( a \sim U(0.5, 2.5), b \sim U(-3, 3), \) and \( c \sim U(0, 0.2). \) This was called the random item pool.

**Error Rate Conditions.** Type I or \( \alpha \) rates were .01 (A1), .05 (A2), and .10 (A3). Type II or \( \beta \) rates were also .01 (B1), .05 (B2), and .10 (B3).

**Hypotheses.** In a mastery testing situation, the usual practice is to establish a single cutoff point along the ability scale, \( \theta_c \), which corresponds to a minimum proportion of items in the domain, \( \pi(\theta_c) \), that an examinee is expected to answer correctly in order to be classified as a master. The relationship between \( \theta_c \) and \( \pi(\theta_c) \), for example, might be

\[
\frac{1}{n} \sum_{j=1}^{n} P_j(\theta_c) = \pi(\theta_c),
\]

where \( n \) is the number of items in the pool representing this testing domain. Because the SPRT procedure requires the setting of two values of \( \theta \) in a simple hypothesis configuration, one usually sets \( \theta_0 < \theta_c < \theta_1 \). The region between \( \theta_0 \) and \( \theta_1 \) is referred to as an indifference region. Reckase (1983) stated that "in order to use the SPRT, a region must be specified around \( \theta_c \) for which it does not matter whether a pass or a fail decision is made. If high accuracy is desired for the decision rule, a narrow indifference region must be specified, but more items will be required to make the decision. As the region gets wider, the decision accuracy declines, but fewer items are required" (p. 243).

In the present study, four simple hypotheses were used to establish four sizes of indifference regions around the chosen value of \( \theta_c = .00 \). These sets of hypotheses \((\theta_0, \theta_1)\) were (1) \( H_1: (-.25, .25) \), (2) \( H_2: (-.5, .5) \), (3) \( H_3: (-.75, .75) \), and (4) \( H_4: (-1.0, 1.0) \).
Results

The results of these 288 computer simulations focused on the effects of the E2 condition on four characteristics or measures of an SPRT: actual or observed α rate ($\alpha^*$), actual or observed β rate ($\beta^*$), average sample number or ASN when $\theta = \theta_0$, and ASN when $\theta = \theta_1$. These results are given in Tables 1 through 6 in terms of overall and marginal means and standard deviations of these variables under the E1 and E2 conditions.

Actual Error Rates

Table 1 shows that even though a nominal type I error or α rate was established prior to the usual SPRT, the observed rate ($\alpha^*$) was actually lower than the nominal one. Under the E1 condition, $\alpha^*$ was .007, .034, and .060, for A1, A2, and A3 nominal rates, respectively. Under the E2 condition, these observed α rates were lower still, .005, .030, and .065, for A1, A2, and A3. However, the overall decrease in $\alpha^*$ for E2 (i.e., from .036 to .033) was quite small and probably insignificant from a practical standpoint.

There was a relatively large decrease in overall mean $\alpha^*$ under E2 for the fourth hypothesis, H4, where the mean $\alpha^* = .027$ (see Table 1). A further analysis of $\alpha^*$ by the nominal error rates, A1, A2, and A3 for this E2-H4 combination revealed that all three values of $\alpha^*$ were lower for H4, although these values were usually lower for each hypothesis under E2, regardless of the nominal α level.

The two exceptions, as seen in Table 2, are at the A3 level. No reasons for these lower $\alpha^*$ were apparent from inspection of further analyses within the design.
Table 3 shows that the observed $B$ rates ($B^*$) were affected even less under the $E_2$ condition than the $a^*$ rates. Although $B^*$ was usually smaller under $E_2$ versus $E_1$, this difference was never greater than .002. However, there was a relatively large decrease in $B^*$ under the $I_4$ condition for both $E_1$ and $E_2$. Table 4 shows that the $B$ rate was lower under all nominal $B$ rates when the item pool consisted of items with variable item parameter values (either known or estimated).

Average Sample Numbers

The overall effect of $E_2$ on average sample number (ASN) was to increase the number of test items required to make a classification decision at each $\theta$ level for which the ASN was analyzed. Table 5 shows that when $\theta = \theta_1$, this overall increase in ASN amounted to 1.1 items from $E_1$ to $E_2$. The greatest increase occurred under the $H_1$ condition (42.5 to 46.8).

Table 6 shows that when $\theta = \theta_0$, the increase in ASN from $E_1$ to $E_2$ was even smaller (.8). Again, the greatest increase occurred under the $H_1$ condition (41.5 to 44.2).

It was interesting to note the effects of different item pools on the ASN. Tables 5 and 6 show that, regardless of the estimation error condition, the ASN increased when items within the pool included a nonzero value for $c$, the pseudo-guessing parameter. When items became more discriminating (i.e., when the discrimination or $a$ parameter changed from 1.0 to 1.5), a decrease in ASN was noted. However, when items had variable item parameters, as was the case under the $I_4$ or random item pool condition, the ASN increased significantly. The observed effects on the ASN under the fixed item pools, $I_1$, $I_2$, and $I_3$, are more easily understood when the hypotheses and the indifference regions are transformed into functions of $\theta_0$ and $\theta_1$, namely $\pi(\theta_0)$ and $\pi(\theta_1)$. Because all of the items in these pools are identical,
\[
\pi(\theta_0) = \frac{c + (1-c)}{1 + \exp (-1.7a(\theta_0 - b))} = \pi_0
\]

and

\[
\pi(\theta_1) = \frac{c + (1-c)}{1 + \exp (-1.7a(\theta_1 - b))} = \pi_1
\]

Table 7 shows these transformed hypotheses and indifference region lengths in terms of \(\pi(\theta_0)\) and \(\pi(\theta_1)\). Wald's SPRT theory predicts that the ASN for any value of \(\theta\) will increase as the size of the indifference region decreases. Therefore, it is no surprise that, of the three fixed pools, the \(I_2\) pool produced the highest ASN at \(\theta_0\) and \(\theta_1\) while \(I_3\) showed the smallest overall ASN values. For the random item pool, \(\pi_0\) and \(\pi_1\) in Table 7 were defined in terms of the averages, \(\bar{\pi}_0\) and \(\bar{\pi}_1\), across the 500 sets of item parameters in \(I_4\), or

\[
\bar{\pi}_0 = \frac{1}{500} \sum_{j=1}^{500} c_j + (1 - c_j)/[1 + \exp(-1.7a_j (\theta_0 - b_j))]
\]

and

\[
\bar{\pi}_1 = \frac{1}{500} \sum_{j=1}^{500} c_j + (1 - c_j)/[1 + \exp(-1.7a_j (\theta_1 - b_j))]
\]

The smaller average indifference regions encountered for \(I_4\) would appear to account for larger ASN values for \(I_4\) in Tables 5 and 6.

Other changes in ASN under the various error rate and hypothesis conditions were again predicted by Wald's SPRT theory. For example, ASN is expected to decrease as \(a\) or \(b\) increases and as the indifference region around \(\theta_c\) increases. Tables 5 and 6 show that this did occur under \(E_1\) and \(E_2\).
Summary and Conclusions

Administering a test using Wald's sequential probability ratio testing procedure on item pools which contain IRT parameter estimates rather than known values did not appear to have much effect on observed mastery or nonmastery classification error rates. These observed error rates were smaller when it was assumed that the item parameters were actually MLEs based on prior calibrations involving examinees with known abilities. However, these smaller observed error rates were not appreciably different from the absent-error condition, E1. Observed error rates under both estimation error conditions were still smaller than the nominal rates established prior to testing and this would appear to be the most important finding regarding error rates.

It should be pointed out that the amount of error in the item parameters was based on several assumptions. First, it was assumed that, during the item calibrations, ability was known. This is rarely true because ability almost always must be estimated in practice. Estimation of ability would increase the amount of error in the item parameter estimates, thereby magnifying the effects of estimation on the SPRT results. Second, the errors were derived under the assumption of normality for the (unidimensional) ability distribution. And finally these error estimates were based on asymptotic standard error formulae and large sample sizes of items and examinees were assumed.

The estimation error condition did appear to have some effect on the observed α rate when the largest indifference region was simulated (H4). How important this effect is in practice remains to be seen because the simulations still produced an α* rate less than the nominal average and because this α* rate occurred with an indifference region (-1.0, 1.0) which may be too large to be useful in actual SPRT administrations.
One noticeable finding involving $\beta^*$ was the amount of decrease in this error rate, regardless of the estimation error condition, when the nature of the item pool changed in terms of item parameters. Wald's SPRT theory makes use of the local independence assumption of IRT through the formulation of the likelihood functions under $H_0$ and $H_1$ as products of probabilities. There is nothing in the SPRT theory which requires that these probabilities be constant from item to item within the pool. And yet, from Table 3, it is obvious that when these probabilities varied considerably from item to item (I4), $\beta^*$ was significantly smaller than when the items did not vary at all (I1, I2 and I3 under E1) or varied by a very small amount (I1, I2, and I3 under E2). A similar effect on $\alpha^*$ was not observed.

On the other hand, the ASN was much larger under the I4 item pool condition, thereby leading to the following conclusion. When items are administered via SPRT procedures and those items vary considerably in $P_i$ for a given examinee, then the ASN will be larger and the $\beta^*$ rate smaller than for SPRT item pools in which the variability of $P_i$ is smaller.

The estimation error condition did yield higher ASN values at all true $\theta$ values, in general, but these increases did not appear to be significant with the item parameter estimation error used in these simulations. According to SPRT theory, the ASN of any SPRT will be a maximum for some $\theta$ value within the indifference region, $(\theta_0, \theta_1)$. The rather large values of ASN for the $H_1$ condition, regardless of estimation error, suggest that this hypothesis could yield ASN values greater than 50 items for some examinees with $\theta$ between -.25 and .25. Therefore, $H_1$ may be an impractical hypothesis to consider for actual SPRT administrations due to the increased test length. Hypothesis $H_2$ or $H_3$ may be more reasonable in practice.

When items from item pools are chosen on some nonrandom basis (e.g., selecting items which maximize $I(\hat{\theta}_n)$ on the basis of estimates of ability, $\hat{\theta}_n$), the
variability of $P_i$ for a given examinee may be minimal, and the effects of using SPRT in a CAT situation, for example, are not expected to change the characteristics of the test from those predicted by the SPRT theory, even when item parameter estimates are used. In fact, when administered as an SPRT, the CAT may even require fewer items and yield smaller classification errors when items are selected for administration on the basis of maximum information. Therefore, a second phase of this research will examine the characteristics of an SPRT when items are administered randomly from $I^4$ versus when the items are administered on the basis of $\max I(\theta)$, with $\theta$ known. A third study will compare the results of the $\max I(\theta)$ procedure of item selection versus a $\max I(\hat{\theta}_n)$ procedure, where $\theta$ is unknown and must be estimated after each item is presented.
REFERENCES


**TABLE 1**

Actual Alpha Rate (α*)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Estimation Error</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>E1</td>
<td>E2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absent</td>
<td>Present</td>
</tr>
<tr>
<td>Overall Mean</td>
<td>144</td>
<td>.036 (0.26)</td>
<td>.033</td>
<td>(.027)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.034 (.026)</td>
<td>.031</td>
<td>(.027)</td>
</tr>
<tr>
<td>Item</td>
<td>36</td>
<td>.039 (.028)</td>
<td>.036</td>
<td>(.027)</td>
</tr>
<tr>
<td>Pool Means</td>
<td>36</td>
<td>.033 (.026)</td>
<td>.033</td>
<td>(.028)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.037 (.027)</td>
<td>.033</td>
<td>(.026)</td>
</tr>
<tr>
<td>α Rate Means</td>
<td>48</td>
<td>.007 (.002)</td>
<td>.005</td>
<td>(.002)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>.034 (.008)</td>
<td>.030</td>
<td>(.009)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>.067 (.014)</td>
<td>.065</td>
<td>(.015)</td>
</tr>
<tr>
<td>α Rate Means</td>
<td>48</td>
<td>.036 (.027)</td>
<td>.033</td>
<td>(.027)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>.036 (.027)</td>
<td>.033</td>
<td>(.027)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>.036 (.026)</td>
<td>.034</td>
<td>(.027)</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>36</td>
<td>.039 (.028)</td>
<td>.037</td>
<td>(.029)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.039 (.027)</td>
<td>.038</td>
<td>(.027)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.032 (.025)</td>
<td>.032</td>
<td>(.027)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.034 (.027)</td>
<td>.027</td>
<td>(.023)</td>
</tr>
</tbody>
</table>

**Note:** Standard deviations are given in parentheses in columns 6 and 8.
TABLE 2

Actual Alpha Rate ($\alpha^*$) Means and Standard Deviations by Hypothesis

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>$N$</th>
<th>$\alpha$</th>
<th>$E_1$ (Estimation Error)</th>
<th>$E_2$ (Estimation Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12</td>
<td>A1</td>
<td>.007 (.002)</td>
<td>.004 (.001)</td>
</tr>
<tr>
<td>$H_1$</td>
<td>12</td>
<td>A2</td>
<td>.038 (.007)</td>
<td>.035 (.007)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A3</td>
<td>.073 (.006)</td>
<td>.072 (.007)</td>
</tr>
<tr>
<td>$H_2$</td>
<td>12</td>
<td>A1</td>
<td>.008 (.002)</td>
<td>.007 (.001)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A2</td>
<td>.038 (.006)</td>
<td>.035 (.008)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A3</td>
<td>.070 (.009)</td>
<td>.071 (.008)</td>
</tr>
<tr>
<td>$H_3$</td>
<td>12</td>
<td>A1</td>
<td>.005 (.002)</td>
<td>.004 (.001)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A2</td>
<td>.029 (.006)</td>
<td>.027 (.008)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A3</td>
<td>.061 (.014)</td>
<td>.065 (.015)</td>
</tr>
<tr>
<td>$H_4$</td>
<td>12</td>
<td>A1</td>
<td>.006 (.003)</td>
<td>.004 (.002)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A2</td>
<td>.032 (.009)</td>
<td>.024 (.006)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A3</td>
<td>.063 (.021)</td>
<td>.052 (.019)</td>
</tr>
</tbody>
</table>

Note: $A_1 = .01$, $A_2 = .05$, and $A_3 = .10$. 
### TABLE 3

**Actual Beta Rate ($\hat{\beta}$)**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>$E_1$ Absent</th>
<th>$E_2$ Present</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Mean</td>
<td>144</td>
<td>.032 (.025)</td>
<td>.031 (.026)</td>
</tr>
<tr>
<td>Item Pool Means</td>
<td>36</td>
<td>.032 (.025)</td>
<td>.035 (.027)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.037 (.027)</td>
<td>.035 (.028)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.032 (.025)</td>
<td>.033 (.028)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.023 (.020)</td>
<td>.022 (.021)</td>
</tr>
<tr>
<td>A Rate Means</td>
<td>48</td>
<td>.032 (.025)</td>
<td>.030 (.026)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>.032 (.025)</td>
<td>.032 (.027)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>.032 (.026)</td>
<td>.031 (.027)</td>
</tr>
<tr>
<td>B Rate Means</td>
<td>48</td>
<td>.007 (.003)</td>
<td>.006 (.002)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>.030 (.011)</td>
<td>.028 (.012)</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>.060 (.019)</td>
<td>.060 (.021)</td>
</tr>
<tr>
<td>Hypothesis Means</td>
<td>36</td>
<td>.041 (.027)</td>
<td>.039 (.030)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.036 (.028)</td>
<td>.034 (.026)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.027 (.022)</td>
<td>.027 (.023)</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>.024 (.020)</td>
<td>.025 (.023)</td>
</tr>
</tbody>
</table>

**Note:** Standard deviations are given in parentheses in columns 6 and 8.
**TABLE 4**

*Actual Beta Rate (β*) Means and Standard Deviations by Item Pool*

<table>
<thead>
<tr>
<th>Item Pool</th>
<th>N</th>
<th>β</th>
<th>Estimation Error</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>E1 Absent</td>
<td>E2 Present</td>
<td></td>
</tr>
<tr>
<td>I1</td>
<td>12</td>
<td>B1</td>
<td>.007 (.002)</td>
<td>.008 (.003)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>B2</td>
<td>.034 (.010)</td>
<td>.033 (.012)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>B3</td>
<td>.066 (.016)</td>
<td>.066 (.018)</td>
<td></td>
</tr>
<tr>
<td>I2</td>
<td>12</td>
<td>B1</td>
<td>.007 (.001)</td>
<td>.006 (.002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>B2</td>
<td>.037 (.005)</td>
<td>.033 (.004)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>B3</td>
<td>.069 (.014)</td>
<td>.066 (.022)</td>
<td></td>
</tr>
<tr>
<td>I3</td>
<td>12</td>
<td>B1</td>
<td>.008 (.002)</td>
<td>.005 (.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>B2</td>
<td>.027 (.012)</td>
<td>.028 (.011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>B3</td>
<td>.061 (.016)</td>
<td>.066 (.014)</td>
<td></td>
</tr>
<tr>
<td>I4</td>
<td>12</td>
<td>B1</td>
<td>.006 (.005)</td>
<td>.004 (.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>B2</td>
<td>.020 (.011)</td>
<td>.019 (.011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>B3</td>
<td>.043 (.019)</td>
<td>.043 (.019)</td>
<td></td>
</tr>
</tbody>
</table>

Note: B1 = .01, B2 = .05, and B3 = .10.
### TABLE 5

**ASN ($H_1$)**

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th></th>
<th>$E1$</th>
<th>$E2$</th>
<th>$E3$</th>
<th>$E4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Absent</td>
<td>Present</td>
</tr>
<tr>
<td>Overall</td>
<td>144</td>
<td></td>
<td>17.6 (19.6)</td>
<td>18.7 (20.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>36</td>
<td>$I_1$</td>
<td>13.5 (14.3)</td>
<td>13.8 (14.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pool</td>
<td>36</td>
<td>$I_2$</td>
<td>16.7 (16.8)</td>
<td>20.0 (20.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td>36</td>
<td>$I_3$</td>
<td>10.2 (9.6)</td>
<td>10.4 (9.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>$I_4$</td>
<td>30.0 (27.6)</td>
<td>30.5 (28.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>$A_1$ (0.01)</td>
<td>22.8 (25.4)</td>
<td>25.5 (27.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate</td>
<td>48</td>
<td>$A_2$ (0.05)</td>
<td>16.9 (17.2)</td>
<td>17.1 (17.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td>48</td>
<td>$A_3$ (0.60)</td>
<td>13.1 (13.4)</td>
<td>13.4 (13.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>$B_1$ (0.01)</td>
<td>18.4 (20.6)</td>
<td>20.0 (22.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate</td>
<td>48</td>
<td>$B_2$ (0.05)</td>
<td>17.1 (19.1)</td>
<td>19.0 (21.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td>48</td>
<td>$B_3$ (0.10)</td>
<td>17.3 (19.4)</td>
<td>17.0 (18.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>$H_1$ (±.25)</td>
<td>42.5 (24.2)</td>
<td>46.8 (24.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothesis</td>
<td>36</td>
<td>$H_2$ (±.50)</td>
<td>14.4 (7.2)</td>
<td>14.3 (7.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td>36</td>
<td>$H_3$ (±.75)</td>
<td>8.2 (3.1)</td>
<td>8.2 (4.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>$H_4$ (±1.90)</td>
<td>5.3 (3.3)</td>
<td>5.5 (3.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Standard deviations are given in parentheses in columns 6 and 8.
<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Estimation Error</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>E1 Absent</td>
<td>E2 Present</td>
<td></td>
</tr>
<tr>
<td>Overall Mean</td>
<td>144</td>
<td>16.2 (19.1)</td>
<td>17.0 (19.7)</td>
<td></td>
</tr>
<tr>
<td>Item Pool Means</td>
<td>36</td>
<td>I1 13.6 (14.6)</td>
<td>13.4 (14.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>I2 16.2 (18.3)</td>
<td>19.3 (20.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>I3 9.4 ( 9.5)</td>
<td>9.4 ( 9.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>I4 25.6 (26.6)</td>
<td>25.9 (26.5)</td>
<td></td>
</tr>
<tr>
<td>α Rate Means</td>
<td>48</td>
<td>A1 (.01) 15.7 (19.1)</td>
<td>18.1 (21.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>A2 (.05) 17.0 (20.1)</td>
<td>17.0 (19.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>A3 (.10) 15.9 (18.6)</td>
<td>15.9 (18.3)</td>
<td></td>
</tr>
<tr>
<td>β Rate Means</td>
<td>48</td>
<td>B1 (.01) 21.8 (25.6)</td>
<td>23.2 (26.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>B2 (.05) 14.6 (15.9)</td>
<td>15.5 (16.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>B3 (.10) 12.2 (12.5)</td>
<td>12.3 (12.7)</td>
<td></td>
</tr>
<tr>
<td>Hypothesis Means</td>
<td>36</td>
<td>H1 (±.75) 41.5 (23.3)</td>
<td>44.2 (22.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>H2 (±.50) 12.4 ( 5.5)</td>
<td>12.8 ( 5.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>H3 (±.75)  6.8 ( 3.1)</td>
<td>6.8 ( 3.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>H4 (±1.00)  4.2 ( 1.7)</td>
<td>4.2 ( 1.8)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard deviations are given in parentheses in columns 6 and 8.


<table>
<thead>
<tr>
<th>Item Pool</th>
<th>Hypothesis</th>
<th>Cutoff Proportions</th>
<th>Indifference Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\pi_0$</td>
<td>$\pi_1$</td>
</tr>
<tr>
<td></td>
<td>H1</td>
<td>.395</td>
<td>.605</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>.299</td>
<td>.701</td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>.218</td>
<td>.782</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>.154</td>
<td>.846</td>
</tr>
<tr>
<td>I1</td>
<td>H1</td>
<td>.516</td>
<td>.684</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>.440</td>
<td>.760</td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>.337</td>
<td>.863</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>.324</td>
<td>.876</td>
</tr>
<tr>
<td>I2</td>
<td>H1</td>
<td>.477</td>
<td>.723</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>.375</td>
<td>.825</td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>.303</td>
<td>.897</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>.258</td>
<td>.942</td>
</tr>
<tr>
<td>I3</td>
<td>H1</td>
<td>.540</td>
<td>.616</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>.503</td>
<td>.655</td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>.466</td>
<td>.692</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>.428</td>
<td>.728</td>
</tr>
</tbody>
</table>

Note: Standard deviations for the indifference regions in I4 are given in parentheses in column 6.
ONR Report Distribution List
May 20, 1988

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
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
Educational Testing Service
Princeton, NJ 08450

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

Dr. Arthur S. Blaisw
Code N711
Naval Training Systems Center
Orlando, FL 32813

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
Section Psychologisch Onderzoek
Rekruterings-En Selectiecentrum
Kwartier Koningin Astrid
Bruininstraat
1120 Brussels, BELGIUM

Dr. Robert Breaux
Code N-095R
Naval Training Systems Center
Orlando, FL 32813

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

Dr. Lyle D. Broemeling
ONR Code 1111SP
800 North Quincy Street
Arlington, VA 22217

Mr. James W. Carey
Commandant (G-PTE)
U.S. Coast Guard
2100 Second Street, S.W.
Washington, DC 20593

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

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

Dr. Robert Carroll
OP 01B7
Washington, DC 20370

Mr. Raymond E. Christal
AFHRL/ME
Brooks AFB, TX 78235
Dr. Benjamin A. Fairbank  
Performance Metrics, Inc.  
5825 Callaghan  
Suite 225  
San Antonio, TX 78228

Dr. Pat Federico  
Code 511  
PNDRC  
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 Program  
P.O. Box 158  
Iowa City, IA 52240

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

Dr. Myron Fischl  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

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  
AFOSH/NL  
Bolling AFB, DC 20332

Dr. Robert D. Gibbons  
Illinois State Psychiatric Inst.  
Rm 529W  
1801 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

Dipl. Pad. Michael W. Habon  
Universitat Dusseldorf  
Erziehungswissenschaftliches Universitätsstr. 1  
D-4000 Dusseldorf 1  
WEST GERMANY

Dr. Ronald K. Hambleton  
Prof. of Education & Psychology  
University of Massachusetts at Amherst  
Hills House  
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 Hatter  
Navy Personnel R&D Center  
Code 62  
San Diego, CA 92152-6800

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

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

Dr. Charles Lewis  
Educational Testing Service  
Princeton, NJ 08541

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

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. Milton Maier  
Center for Naval Analysis  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0268

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

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

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

Dr. James McBride  
Psychological Corporation  
c/o Harcourt, Brace, Javanovich Inc.  
1250 West 6th Street  
San Diego, CA 92101

Dr. Clarence McCormick  
HQ, MEPCOM  
MEPCT-P  
2500 Green Bay Road  
North Chicago, IL 60064

Dr. Robert McKinley  
Educational Testing Service  
20-P  
Princeton, NJ 08541

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

Dr. Barbara Means  
Human Resources Research Organization  
1100 South Washington  
Alexandria, VA 22314

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 20330

Dr. W. Alan Nicewander  
University of Oklahoma  
Department of Psychology  
Oklahoma City, OK 73103
Dr. John Tangney
AFOSR/NL
Bolling AFB, DC 20332

Dr. Kikumi Tatsuoka
CERL
252 Engineering Research Laboratory
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
Champaign, IL 61820

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 310
St. Paul, MN 55114

Dr. Frank Vicingo
Naval Personnel R&D Center
San Diego, CA 92152-6800

Dr. Howard Wainer
Division of Psychological Studies
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
Program Manager
Manpower Analysis Program
HumRRO
1100 S. Washington St.
Alexandria, VA 22314

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

Dr. Ronald A. Weitzman
NPS, Code 54Wz
Monterey, CA 92215-6800

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Douglas Wetzel
Code 12
Naval Personnel R&D Center
San Diego, CA 92152-6800

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