

**ACT Research Report Series**

**87-17**

**AD-A196 674**

**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**

Prepared under Contract No. N00014-85-C-0241, Contract Authority Identification No. NR 154-531, with the Cognitive Science Research Program of the Office of Naval Research.

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

**September 1987**  
*October*

**DTIC ELECTE**  
**S** **D**  
JUL 07 1988  
H

**ACT**

**Best Available Copy**

**2004 0323 024**

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release: distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) <b>ONR 87-1</b>		5. MONITORING ORGANIZATION REPORT NUMBER(S)	
6a. NAME OF PERFORMING ORGANIZATION <b>ACT</b>	6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION <b>COGNITIVE SCIENCE RESEARCH PROGRAMS OFFICE OF NAVAL RESEARCH</b>	
6c. ADDRESS (City, State, and ZIP Code) <b>P.O. Box 168 Iowa City, IA 52243</b>		7b. ADDRESS (City, State, and ZIP Code) <b>Code 1142CS Arlington, VA 22217-5000</b>	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER <b>N00014-85-C-0241</b>	
8c. ADDRESS (City, State, and ZIP Code)		10. SOURCE OF FUNDING NUMBERS	
		PROGRAM ELEMENT NO. <b>61153N</b>	PROJECT NO. <b>RR04204</b>
11. TITLE (Include Security Classification) <b>The effect of item parameter estimation error on decisions made using the sequential probability ratio test</b>			
12. PERSONAL AUTHOR(S) <b>Judith A. Spray, Mark D. Reckase</b>			
13a. TYPE OF REPORT <b>Technical</b>	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Year, Month, Day) <b>1987, October</b>	15. PAGE COUNT <b>26</b>
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) <b>Adaptive testing      Sequential probability ratio tests Item response theory      alpha      beta Latent trait theory</b>	
FIELD <b>05</b>	GROUP <b>09</b>		
19. ABSTRACT (Continue on reverse if necessary and identify by block number) 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 <u>a</u> -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 <u>a</u> , <u>b</u> and <u>c</u> 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. <i>Keywords: adaptive testing, (KP)</i>			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>	
22a. NAME OF RESPONSIBLE INDIVIDUAL <b>Dr. Charles Davis</b>		22b. TELEPHONE (Include Area Code) <b>(202) 696-4046</b>	22c. OFFICE SYMBOL <b>ONR 1142CS</b>

**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.



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

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  $\underline{X} = \underline{x}$  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  $\underline{x}$ . A ratio of these two functions,  $L(\underline{x}) = \pi(\theta_1)/\pi(\theta_0)$ , is called a likelihood ratio.

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

$$\text{Prob (choosing } H_1 | H_0 \text{ is true)} = \alpha$$

and

$$\text{Prob (choosing } H_0 | H_1 \text{ is true)} = \beta .$$

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

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

and

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

According to Wald's SPRT, trials or items would be observed in sequence,  $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ , and following each observation, the likelihood ratio,  $L(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$ , would be computed, where

$$L(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) = \frac{\pi_1(\theta_1) \cdot \pi_2(\theta_1) \cdots \pi_n(\theta_1)}{\pi_1(\theta_0) \cdot \pi_2(\theta_0) \cdots \pi_n(\theta_0)} .$$

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

$$L(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) \geq A,$$

then  $H_1$  is accepted. If

$$L(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) \leq B,$$

then  $H_0$  is accepted. If

$$B < L(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) < A,$$

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

Once  $\alpha$ ,  $\beta$  and the hypotheses are set prior to testing, the stopping rules of the test (i.e., the boundaries) are defined. Although  $\alpha$  and  $\beta$  are determined prior to observing  $\underline{x}$ , where  $\underline{x} = (\underline{x}_1 \underline{x}_2 \dots \underline{x}_n)$ , Wald (1947) pointed out that the actual error rates observed in practice,  $\alpha^*$  and  $\beta^*$ , would be bounded from above by

$$\alpha^* \leq \alpha/(1-\beta)$$

and

$$\beta^* \leq \beta/(1-\alpha)$$

(see Wald, 1947, p. 46). This means that even though the nominal error probabilities,  $\alpha$  and  $\beta$ , 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

$$I(\hat{\theta}_n) = \frac{P'(\hat{\theta}_n)}{P(\hat{\theta}_n)Q(\hat{\theta}_n)},$$

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  $L(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$  is written as the product of the individual item likelihood ratios through the assumption of local independence of the  $\underline{x}_i$ , given  $\theta_i$ .

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  $\max I(\hat{\theta}_n)$ , 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,  $\pi(\theta_c)$ .

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  $L(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$  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  $\alpha$ ,  $\beta$ ,  $\theta_0$  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 X 4 X 3 X 3 X 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 ( $\alpha$ ), (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 condition, 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  $\ell$  is this sum, or

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

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

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

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

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

where  $\phi(\theta)$  is taken to be a normal density with zero mean and variance one (Thissen & Wainer, 1982). In other words, if  $\Sigma$  is the variance-covariance matrix of  $\xi$ , then  $\Sigma$  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,

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

For the first pool (I1),  $\underline{a} = 1$ ,  $\underline{b} = 0$ , and  $\underline{c} = 0$  for all 500 items. Under the E1 condition, these identical items represented a simple SPRT with constant success probability,  $P(\theta_i)$  for a given  $\theta_i$  value. Under the E2 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  $\underline{a}_1 = \underline{a}_2 = \dots = \underline{a}_{500}$ , each  $\underline{a}$  parameter represented an estimate,  $\hat{\underline{a}}_j$ , where

$$\hat{\underline{a}}_j = \underline{a} + \epsilon_{\underline{a}j},$$

and  $\epsilon_{\underline{a}j}$  was a random deviate from a multivariate normal distribution with mean vector 0 and variance-covariance matrix  $\Sigma$ , defined previously.

For the second item pool (I2),  $\underline{a} = 1$ ,  $\underline{b} = 0$ , and  $\underline{c} = .2$ . For the third pool (I3),  $\underline{a} = 1.5$ ,  $\underline{b} = 0$ , and  $\underline{c} = .2$ . Again, under E1 these item parameters remained constant for all 500 items in a pool. However, under E2, item parameter values were assumed to be estimates ( $\underline{a} + \epsilon_{\underline{a}j}$ ,  $\underline{b} + \epsilon_{\underline{b}j}$ , and  $\underline{c} + \epsilon_{\underline{c}j}$  with  $\epsilon_{\underline{a}j}$ ,  $\epsilon_{\underline{b}j}$ , and  $\epsilon_{\underline{c}j}$  being random deviates as before).

For the fourth item pool (I4), the 500 sets of parameters were generated from a pseudo-random number generator with  $a \sim U(.5, 2.5)$ ,  $b \sim U(-3., 3.)$ , and  $c \sim U(.0, .2)$ . 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) H1:  $(-.25, .25)$ , (2) H2:  $(-.5, .5)$ , (3) H3:  $(-.75, .75)$ , and (4) H4:  $(-1.0, 1.0)$ .

## 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  $\alpha$  rate ( $\alpha^*$ ), actual or observed  $\beta$  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  $\alpha$  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  $\alpha$  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  $\alpha$  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  $\beta$  rates ( $\beta^*$ ) were affected even less under the E2 condition than the  $\alpha^*$  rates. Although  $\beta^*$  was usually smaller under E2 versus E1, this difference was never greater than .002. However, there was a relatively large decrease in  $\beta^*$  under the I4 condition for both E1 and E2. Table 4 shows that the  $\beta^*$  rate was lower under all nominal  $\beta$  rates when the item pool consisted of items with variable item parameter values (either known or estimated).

#### Average Sample Numbers

The overall effect of E2 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 E1 to E2. The greatest increase occurred under the H1 condition (42.5 to 46.8).

Table 6 shows that when  $\theta = \theta_0$ , the increase in ASN from E1 to E2 was even smaller (.8). Again, the greatest increase occurred under the H1 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 I4 or random item pool condition, the ASN increased significantly. The observed effects on the ASN under the fixed item pools, I1, I2, and I3, 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 I2 pool produced the highest ASN at  $\theta_0$  and  $\theta_1$  while I3 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 I4, or

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

and

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

The smaller average indifference regions encountered for I4 would appear to account for larger ASN values for I4 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  $\alpha$  or  $\beta$  increases and as the indifference region around  $\theta_c$  increases.

Tables 5 and 6 show that this did occur under E1 and E2.

### 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,  $E_1$ . 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  $\alpha$  rate when the largest indifference region was simulated ( $H_4$ ). How important this effect is in practice remains to be seen because the simulations still produced an  $\alpha^*$  rate less than the nominal average and because this  $\alpha^*$  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 H1 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, H1 may be an impractical hypothesis to consider for actual SPRT administrations due to the increased test length. Hypothesis H2 or H3 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

- Reckase, M. D. (1983). A procedure for decision making using tailored testing. In D. J. Weiss (Ed.), New horizons in testing: Latent trait test theory and computerized adaptive testing (pp. 237-255). New York: Academic Press.
- Thissen, D., & Wainer, H. (1982). Some standard errors in item response theory. Psychometrika, 47, 397-412.
- Wald, A. (1947). Sequential analysis. New York: Wiley.

TABLE 1  
Actual Alpha Rate ( $\alpha^*$ )

	<u>N</u>			Estimation Error	
				<u>E1</u> Absent	<u>E2</u> Present
<u>Overall Mean</u>	144			.036 (0.26)	.033 (.027)
	36	<u>I1</u>		.034 (.026)	.031 (.027)
<u>Item Pool Means</u>	36	<u>I2</u>		.039 (.028)	.036 (.027)
	36	<u>I3</u>		.033 (.026)	.033 (.028)
	36	<u>I4</u>		.037 (.027)	.033 (.026)
<u><math>\alpha</math> Rate Means</u>	48	<u>A1</u>	(.01)	.007 (.002)	.005 (.002)
	48	<u>A2</u>	(.05)	.034 (.008)	.030 (.009)
	48	<u>A3</u>	(.10)	.067 (.014)	.065 (.015)
<u><math>\beta</math> Rate Means</u>	48	<u>B1</u>	(.01)	.036 (.027)	.033 (.027)
	48	<u>B2</u>	(.05)	.036 (.027)	.033 (.027)
	48	<u>B3</u>	(.10)	.036 (.026)	.034 (.027)
<u>Hypothesis Means</u>	36	<u>H1</u>	( $\pm .25$ )	.039 (.028)	.037 (.029)
	36	<u>H2</u>	( $\pm .50$ )	.039 (.027)	.038 (.027)
	36	<u>H3</u>	( $\pm .75$ )	.032 (.025)	.032 (.027)
	36	<u>H4</u>	( $\pm 1.00$ )	.034 (.027)	.027 (.023)

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

TABLE 2

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

	<u>N</u>	<u><math>\alpha</math></u>	Estimation Error	
			<u>E1</u> Absent	<u>E2</u> Present
<u>H1</u>	12	<u>A1</u>	.007 (.002)	.004 (.001)
	12	<u>A2</u>	.038 (.007)	.035 (.007)
	12	<u>A3</u>	.073 (.006)	.072 (.007)
<u>H2</u>	12	<u>A1</u>	.008 (.002)	.007 (.001)
	12	<u>A2</u>	.038 (.006)	.035 (.008)
	12	<u>A3</u>	.070 (.009)	.071 (.008)
<u>H3</u>	12	<u>A1</u>	.005 (.002)	.004 (.001)
	12	<u>A2</u>	.029 (.006)	.027 (.008)
	12	<u>A3</u>	.061 (.014)	.065 (.015)
<u>H4</u>	12	<u>A1</u>	.006 (.003)	.004 (.002)
	12	<u>A2</u>	.032 (.009)	.024 (.006)
	12	<u>A3</u>	.063 (.021)	.052 (.019)

Note: A1 = .01, A2 = .05, and A3 = .10.

TABLE 3  
Actual Beta Rate ( $\beta^*$ )

	<u>N</u>	Estimation Error		
		<u>E1</u> Absent	<u>E2</u> Present	
<u>Overall Mean</u>	144	.032 (.025)	.031 (.026)	
<u>Item Pool Means</u>	36	<u>I1</u>	.036 (.027)	.035 (.027)
	36	<u>I2</u>	.037 (.027)	.035 (.028)
	36	<u>I3</u>	.032 (.025)	.033 (.028)
	36	<u>I4</u>	.023 (.020)	.022 (.021)
<u><math>\alpha</math> Rate Means</u>	48	<u>A1</u> (.01)	.032 (.025)	.030 (.026)
	48	<u>A2</u> (.05)	.032 (.025)	.032 (.027)
	48	<u>A3</u> (.10)	.032 (.026)	.031 (.027)
<u><math>\beta</math> Rate Means</u>	48	<u>B1</u> (.01)	.007 (.003)	.006 (.002)
	48	<u>B2</u> (.05)	.030 (.011)	.028 (.012)
	48	<u>B3</u> (.10)	.060 (.019)	.060 (.021)
<u>Hypothesis Means</u>	36	<u>H1</u> ( $\pm .25$ )	.041 (.027)	.039 (.030)
	36	<u>H2</u> ( $\pm .50$ )	.036 (.028)	.034 (.026)
	36	<u>H3</u> ( $\pm .75$ )	.027 (.022)	.027 (.023)
	36	<u>H4</u> ( $\pm 1.00$ )	.024 (.020)	.025 (.023)

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

TABLE 4  
Actual Beta Rate ( $\beta$ ) Means and Standard Deviations by Item Pool

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

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

TABLE 5

ASN ( $H_1$ )

	<u>N</u>			Estimation Error	
				<u>E1</u> Absent	<u>E2</u> Present
<u>Overall Mean</u>	144			17.6 (19.6)	18.7 (20.9)
	36	<u>I1</u>		13.5 (14.3)	13.8 (14.7)
<u>Item Pool Means</u>	36	<u>I2</u>		16.7 (16.8)	20.0 (20.5)
	36	<u>I3</u>		10.2 (9.6)	10.4 (9.9)
	36	<u>I4</u>		30.0 (27.6)	30.5 (28.6)
<u><math>\alpha</math> Rate Means</u>	48	<u>A1</u>	(.01)	22.8 (25.4)	25.5 (27.5)
	48	<u>A2</u>	(.05)	16.9 (17.2)	17.1 (17.8)
	48	<u>A3</u>	(.60)	13.1 (13.4)	13.4 (13.8)
<u><math>\beta</math> Rate Means</u>	48	<u>B1</u>	(.01)	18.4 (20.6)	20.0 (22.6)
	48	<u>B2</u>	(.05)	17.1 (19.1)	19.0 (21.7)
	48	<u>B3</u>	(.10)	17.3 (19.4)	17.0 (18.7)
<u>Hypothesis Means</u>	36	<u>H1</u>	( $\pm$ .25)	42.5 (24.2)	46.8 (24.1)
	36	<u>H2</u>	( $\pm$ .50)	14.4 (7.2)	14.3 (7.1)
	36	<u>H3</u>	( $\pm$ .75)	8.2 (5.1)	8.2 (4.9)
	36	<u>H4</u>	( $\pm$ 1.00)	5.3 (3.3)	5.5 (3.3)

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

TABLE 6

ASN( $H_0$ )

	<u>N</u>			Estimation Error	
				<u>E1</u> Absent	<u>E2</u> Present
Overall <u>Mean</u>	144			16.2 (19.1)	17.0 (19.7)
	36	<u>I1</u>		13.6 (14.6)	13.4 (14.0)
Item Pool <u>Means</u>	36	<u>I2</u>		16.2 (18.3)	19.3 (20.9)
	36	<u>I3</u>		9.4 ( 9.5)	9.4 ( 9.4)
	36	<u>I4</u>		25.6 (26.6)	25.9 (26.5)
$\alpha$ Rate <u>Means</u>	48	<u>A1</u>	(.01)	15.7 (19.1)	18.1 (21.2)
	48	<u>A2</u>	(.05)	17.0 (20.1)	17.0 (19.8)
	48	<u>A3</u>	(.10)	15.9 (18.6)	15.9 (18.3)
$\beta$ Rate <u>Means</u>	48	<u>B1</u>	(.01)	21.8 (25.6)	23.2 (26.4)
	48	<u>B2</u>	(.05)	14.6 (15.9)	15.5 (16.2)
	48	<u>B3</u>	(.10)	12.2 (12.5)	12.3 (12.7)
Hypothesis <u>Means</u>	36	<u>H1</u>	( $\pm$ .75)	41.5 (23.3)	44.2 (22.0)
	36	<u>H2</u>	( $\pm$ .50)	12.4 ( 5.5)	12.8 ( 5.9)
	36	<u>H3</u>	( $\pm$ .75)	6.8 ( 3.1)	6.8 ( 3.1)
	36	<u>H4</u>	( $\pm$ 1.00)	4.2 ( 1.7)	4.2 ( 1.8)

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

TABLE 7

Hypotheses and Indifference Regions in Terms of  $\pi(\theta)$ 

Item Pool	Hypothesis	Cutoff Proportions		Indifference Region
		$\pi_0$	$\pi_1$	$(\pi_1 - \pi_0)$
<u>I1</u>	<u>H1</u>	.395	.605	.210
	<u>H2</u>	.299	.701	.402
	<u>H3</u>	.218	.782	.564
	<u>H4</u>	.154	.846	.692
<u>I2</u>	<u>H1</u>	.516	.684	.168
	<u>H2</u>	.440	.760	.320
	<u>H3</u>	.337	.863	.526
	<u>H4</u>	.324	.876	.552
<u>I3</u>	<u>H1</u>	.477	.723	.246
	<u>H2</u>	.375	.825	.450
	<u>H3</u>	.303	.897	.594
	<u>H4</u>	.258	.942	.684
<u>I4</u>	<u>H1</u>	.540	.616	.076 (.093)
	<u>H2</u>	.503	.655	.152 (.172)
	<u>H3</u>	.466	.692	.226 (.230)
	<u>H4</u>	.428	.728	.300 (.270)

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  
Studivstraede 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. Blaiwes  
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  
Sectie Psychologisch Onderzoek  
Rekruterings-En Selectiecentrum  
Kwartier Koningen Astrid  
Bruijnstraat  
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/MOE  
Brooks AFB, TX 78235

Best Available Copy

Dr. Norman Cliff  
Department of Psychology  
Univ. of So. California  
University Park  
Los Angeles, CA 90007

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 Crombag  
University of Leyden  
Education Research Center  
Boerhaavelaan 2  
2334 EN Leyden  
The NETHERLANDS

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. DeAvala  
Measurement, Statistics,  
and Evaluation  
Benjamin Building  
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  
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  
425 Fraser  
Lawrence, KS 66045

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

**BEST AVAILABLE COPY**

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

Dr. Pat Federico  
Code 511  
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  
Program  
P.O. Box 168  
Iowa City, IA 52240

Dr. Gerhard Fischer  
Liebiggasse 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  
AFOSR/NL  
Bolling AFB, DC 20332

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

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 Hetter  
Navy Personnel R&D Center  
Code 62  
San Diego, CA 92152-6800

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

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

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

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  
Department of Education  
University of Alberta  
Edmonton, Alberta  
CANADA

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

Dr. Robert Jannarone  
Department of Psychology  
University of South Carolina  
Columbia, SC 29208

Dr. Dennis E. Jennings  
Department of Statistics  
University of Illinois  
1409 West Green Street  
Urbana, IL 61801

Dr. Douglas H. Jones  
Thatcher Jones Associates  
P.O. Box 6640  
10 Trafalgar Court  
Lawrenceville, NJ 08646

Dr. Milton S. Katz  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

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  
University of Texas-Austin  
Measurement and Evaluation  
Center  
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  
San Diego, CA 92152-6800

Dr. Daryll Lang  
Navy Personnel R&D Center  
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

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 08451

Dr. Clessen Martin  
Army Research Institute  
5001 Eisenhower Blvd.  
Alexandria, VA 22303

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 Mitlev  
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  
Oklahoma City, OK 73069

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

Fleet Support Office,  
NPRDC (Code 301)  
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 Olson  
WICAT, Inc.  
1875 South State Street  
Orem, UT 84057

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 01B7  
Washington, DC 20370

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

Dr. Jesse Orlansky  
Institute for Defense Analyses  
1301 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

Administrative Sciences Department,  
Naval Postgraduate School  
Monterey, CA 93940

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/MP  
Brooks AFB, TX 78235

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

Dr. Carl Ross  
CNET-PDCD  
Building 90  
Great Lakes NTC, IL 60089

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

Dr. Fumiko Samejima  
Department of Psychology  
University of Tennessee  
3108 AustinPeay 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 Shigemasa  
7-9-24 Kugenuma-Kaigan  
Fujusawa 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  
Alexandria, VA 22314

Dr. Richard E. Snow  
Department of Psychology  
Stanford University  
Stanford, CA 94306

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

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 Symson  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

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 Vicino  
Navv 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 92152-6800

Major John Welsh  
AFHRL/MOAN  
Brooks AFB, TX 78223

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

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

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

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

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

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

Dr. Martin F. Wiskoff  
Navy Personnel R & D Center  
San Diego, CA 92152-6800

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  
San Diego, CA 92152-6800

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

Dr. Wendy Yen  
CTS/McGraw Hill  
Del Monte Research Park  
Monterey, CA 95940

Dr. Joseph L. Young  
Memory & Cognitive  
Processes  
National Science Foundation  
Washington, DC 20550