DOCUMENT RESUME

ED 322 212 TM 015 457

AUTHOR	Samejima, Fumiko
TITLE	Validity Measures in the Context of Latent Trait Models.
INSTITUTION	Tennessee Univ., Knoxville. Dept. of Psychology.
SPONS AGENCY	Office of Naval Research, Arlington, Va.
REPORT NO	ONR-RR-90-3
PUB DATE	15 Jun 90
CONTRACT	N00014-87-K-0320-4421-549
NOTE	30p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Adaptive Testing; Computer Assisted Testing; Equations (Mathematics); *Item Response Theory; *Mathematical Models; Test Items; Test Reliability; *Test Validity
IDENTIFIERS	Classical Test Theory; Information Function (Tests); *Theta Estimates

ABSTRACT

Test validity is a concept that has often been ignored in the context of latent trait models and in modern test theory, particularly as it relates to computerized adaptive testing. Some considerations about the validity of a test and of a single item are proposed. This paper focuses on measures that are population-free and that will provide local and abundant information just as the information functions do in comparison with the test reliability coefficient in classical mental test theory. In so doing, validity indices for different purposes of testing and those that are tailored for a specific population of examinees are considered. The resulting indices should not be incidental as those in classical mental test theory are; they are truly attributes of the item and the test. Six figures illustrate the discussion. (Author/SLD)

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; Distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Fumiko Samejima, Ph.D. Psychology Department; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Cognitive Science 1142 CS
6c. ADDRESS (City, State, and ZIP Code): 310B Austin Peay Building, The University of Tennessee, Knoxville, TN 37996-0900	7b. ADDRESS (City, State, and ZIP Code): Office of Naval Research, 800 N. Quincy Street, Arlington, VA 22217
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Cognitive Science Research Program; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-87-K-0320
8c. ADDRESS (City, State, and ZIP Code): Office of Naval Research, 800 N. Quincy Street, Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 61153N; PROJECT NO RR-042-04; TASK NO 042-04-01; WORK UNIT ACCESSION NO 4421-549

11. TITLE (Include Security Classification): Validity measures in the context of latent trait models

12. PERSONAL AUTHOR(S): Fumiko Samejima, Ph.D.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
technical report	FROM 1987 TO 1990	June 15, 1990	26

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Latent Trait Models, Item Validity, Test Validity, Computerized Adaptive Testing

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

In contrast to the progressive desolution of the reliability coefficient in classical mental test theory and the replacement by the test information function in latent trait models, the issue of test validity has been more or less neglected in modern mental test theory. The present paper proposes some considerations about the validity of a test and of a single item. Effort has been focused upon searching for measures which are population-free, and which will provide us with local and abundant information just as the information functions do in comparison with the test reliability coefficient in classical mental test theory. In so doing, validity indices for different purposes of testing and also those which are tailored for a specific population of examinees are consider

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL: Dr. Charles E. Davis	22b. TELEPHONE (Include Area Code): 202-696-4046; 22c. OFFICE SYMBOL: ONR-1142-CS

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

S/N 0102-LF-014-6603

TABLE OF CONTENTS

The research was conducted at the principal investigator's laboratory, 405 Austin Peay Bldg., Department of Psychology, University of Tennessee, Knoxville, Tennessee. Those who worked as assistants for this research include Christine A. Golik, Barbara A. Livingston, Lee Hai Gan and Nancy H. Domm.

I Introduction

In classical mental test theory, the reliability and the validity coefficients of a test are considered to be two essential topics. In modern mental test theory, or in latent trait models, this is not the case, however. In particular, test validity is one concept that has been neglected in the context of latent trait models.

Several types of validity have been identified and discussed in classical mental test theory, which include content validity, construct validity, and criterion-oriented validity. Perhaps we can say that, in modern mental test theory, both content validity and construct validity are well accomodated, although they are not explicitly stated. If each item is based upon cognitive processes that are directly related to the ability to be measured, then the content of the operationally defined latent variable behind the examinees' performances will be validated. Also construct validity can be identified, with all the mathematically sophisticated structures and functions which characterize latent trait models and which classical mental test theory does not provide.

With respect to the criterion-oriented validity, however, so far latent trait models have not offered so much as they did to the test reliability and to the standard error of measurement (cf. Samejima, 1977, 1990). From the scientific point of view, however, we need to confirm if, indeed, the test measures what it is supposed to measure, even if we have chosen our items carefully enough in regard to their contents, and even if we are equipped with highly sophisticated mathematics.

In classical mental test theory, the validity coefficient is a single number, i.e., the product-moment correlation coefficient between the test score and the criterion variable. Researchers tend to put too much faith in the validity coefficient, or in the reliability coefficient, however. The correlation coefficient is largely affected by the heterogeneity of the group of examinees, i.e., for a fixed test the coefficient tends to be higher when individual differences among the examinees in the group are greater, and vice versa (cf. Samejima, 1977). Thus we must keep in mind that *so-called* test validity represents the degree of heterogeneity in ability among the examinees tested, as well as the quality of the test itself.

By virtue of the population-free nature of latent trait theory, we should be able to find some indices of item validity, and of test validity, which are not affected by the group of examinees. The resulting indices should not be *incidental* as those in classical mental test theory are, but truly be attributes of the item and the test themselves.

In the present research an attempt has been made to obtain such population-free measures of item validity and of test validity, which are basically *locally* defined.

II Performance Function: Regression of the External Criterion Variable on the Latent Variable

Let θ be ability, or latent trait, which assumes any real number. We assume that there is a set of n test items measuring θ whose characteristics are known. Let g denote such an item, k_g be a discrete item response to item g, and $P_{k_g}(\theta)$ denote the *operating characteristic* of k_g, or the conditional probability assigned to k_g, given θ, i.e.,

$$P_{k_g}(\theta) = Prob.[k_g \mid \theta] \quad . \tag{2.1}$$

We assume that $P_{k_g}(\theta)$ is three-times differentiable with respect to θ. We have for the *item response information function*

(2.2) $$I_{k_g}(\theta) = -\frac{\partial^2}{\partial\theta^2} \log P_{k_g}(\theta) \quad ,$$

and the *item information function* is defined as the conditional expectation of $I_{k_g}(\theta)$, given θ, so that we can write

(2.3) $$I_g(\theta) = E[I_{k_g}(\theta) \mid \theta] = \sum_{k_g} I_{k_g}(\theta) P_{k_g}(\theta) \quad .$$

In the special case where the item g is scored dichotomously, this item information function is simplified to become

(2.4) $$I_g(\theta) = [\frac{\partial}{\partial\theta} P_g(\theta)]^2 \; [\{P_g(\theta)\}\{1 - P_g(\theta)\}]^{-1} \quad ,$$

where $P_g(\theta)$ is the operating characteristic of the correct answer to item g.

Let V be a *response pattern* such that

(2.5) $$V = \{\ k_g\ \}' \qquad g = 1, 2, ..., n \quad .$$

The operating characteristic, $P_V(\theta)$, of the response patten V is defined as the conditional probability of V, given θ, and by virtue of *local independence* we can write

(2.6) $$P_V(\theta) = \prod_{k_g \epsilon V} P_{k_g}(\theta) \quad .$$

The *response pattern information function,* $I_V(\theta)$, is given by

(2.7) $$I_V(\theta) = -\frac{\partial^2}{\partial\theta^2} \log P_V(\theta) = \sum_{k_g \epsilon V} I_{k_g}(\theta) \quad ,$$

and the *test information function,* $I(\theta)$, is defined as the conditional expectation of $I_V(\theta)$, given θ, and we obtain from (2.2), (2.3), (2.5), (2.6) and (2.7)

(2.8) $$I(\theta) = E[I_V(\theta) \mid \theta] = \sum_V I_V(\theta) P_V(\theta) = \sum_{g=1}^{n} I_g(\theta) \quad .$$

A big advantage of the modern mental test theory is that the standard error of estimation can *locally* be defined by using $[I(\theta)]^{-1/2}$. Unlike in classical mental test theory this function *does not depend upon the population of examinees,* but is solely a property of the test itself, which should be the way if we call it the standard error, or the reliability, *of a test.* It is well known that this function provides us with the asymptotic standard deviation of the conditional distribution of the maximum likelihood estimate of θ, given its true value.

It is assumed that there exists an external criterion variable, which can be measured directly or indirectly. This is the situation which is also assumed when we deal with *criterion-oriented validity* or *predictive validity* in classical mental test theory.

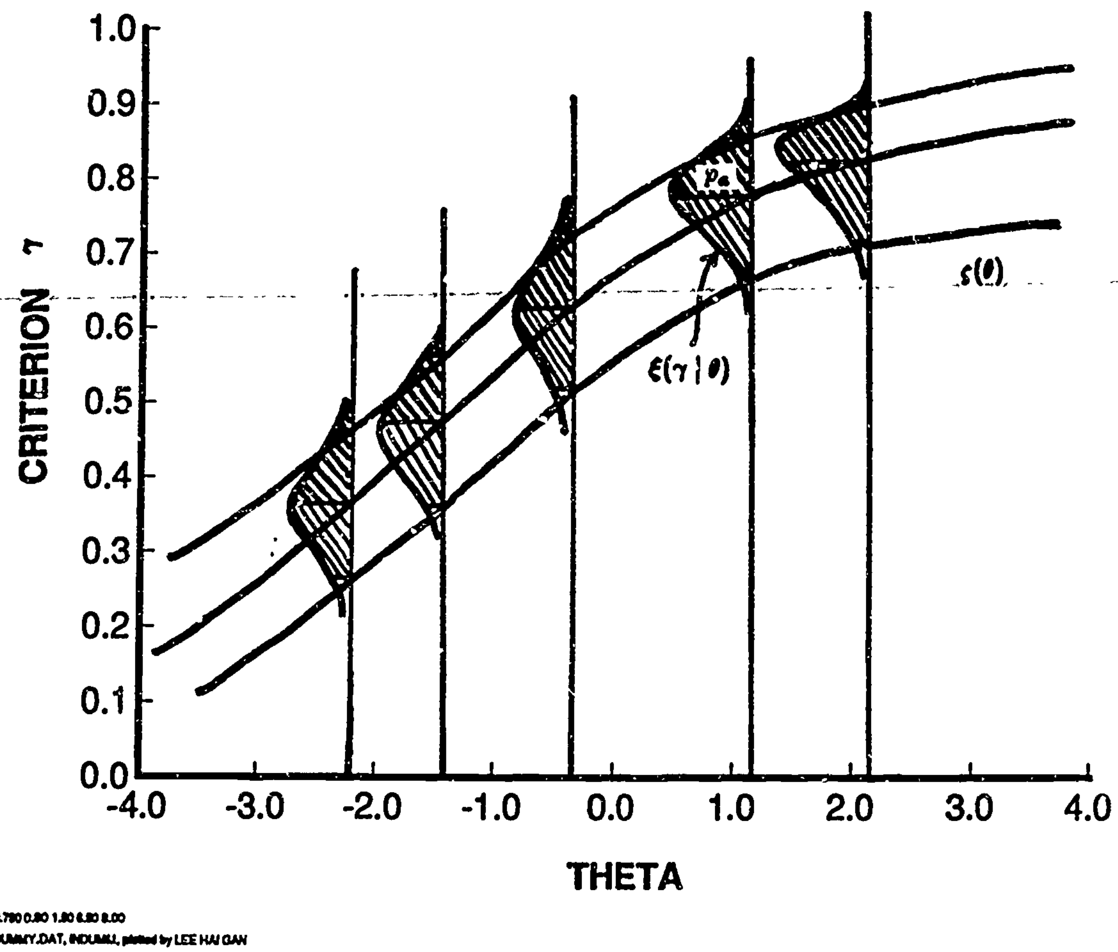

FIGURE 2-1

Relationships among θ , γ , p_a , $\xi(\gamma\,|\,\theta)$ and $\varsigma(\theta)$.

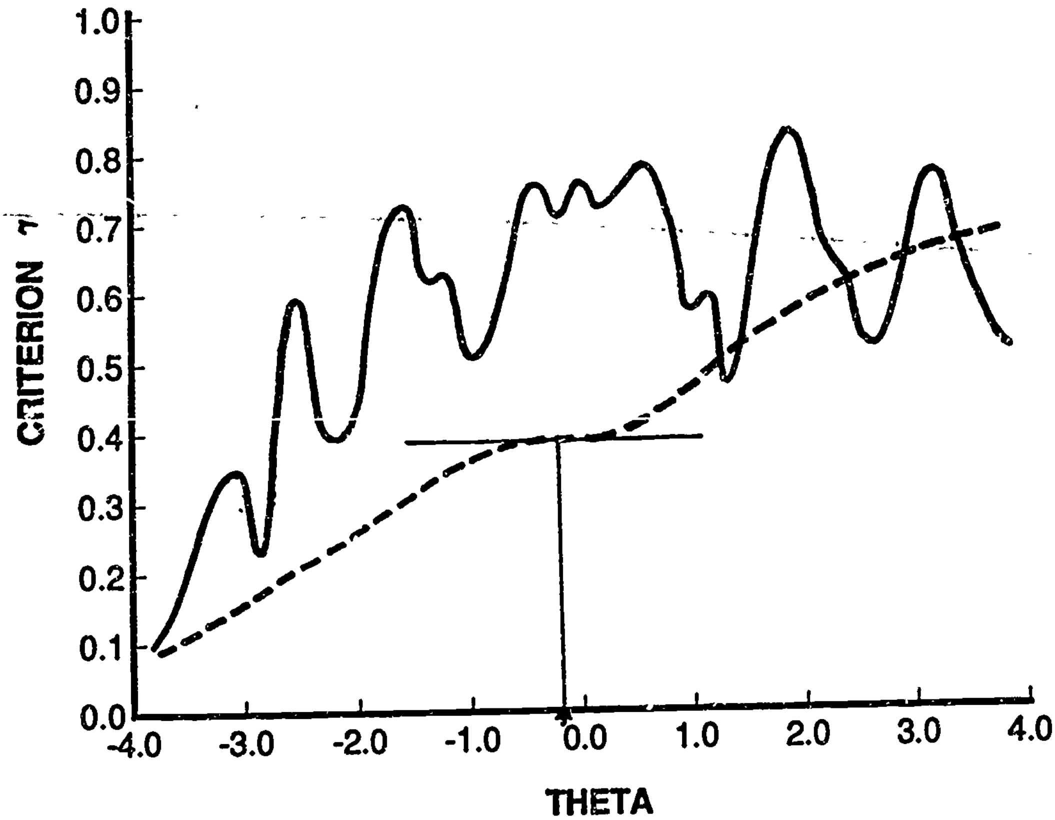

FIGURE 2-2

Two Hypothetical Performance Functions $\varsigma(\theta)$, One of Which Is Not Likely to Be the Case (Solid Line), and the Other Has a Derivative Equal to Zero at One Point of θ (Dashed Line).

Let γ denote the *criterion variable*, representing the performance in a specific job, etc. We shall consider the conditional density of the criterion performance, given ability, and denote it by $\xi(\gamma \mid \theta)$. The *performance function*, $\varsigma(\theta)$, can be defined as the regression of γ on θ, or by taking, say, the 75, 90 or 95 percentile point of each conditional distribution of γ, given θ. Let p_a denote the probability which is large enough to satisfy us as a confidence level. Thus we can write

$$p_a = \int_{\varsigma(\theta)}^{\bar{\gamma}} \xi(\gamma \mid \theta)\, d\gamma \ , \tag{2.9}$$

where $\bar{\gamma}$ denotes the least upper bound of the criterion variable γ.

Figure 2-1 illustrates the relationships among θ, γ, p_a, $\xi(\gamma \mid \theta)$ and $\varsigma(\theta)$. It may be reasonable to assume that the functional relationship between θ and $\varsigma(\theta)$ is relatively simple, not as is illustrated by the solid line in Figure 2-2, i.e., we do not expect $\varsigma(\theta)$ to go up and down frequently within a relatively short range of θ. We shall assume that $\varsigma(\theta)$ is twice differentiable with respect to θ.

In dealing with an additional dimension or dimensions, i.e., the criterion variable or variables, in latent space, one of the most difficult things is to keep the population-free nature, which is characteristic of the latent trait models, the main feature that distinguishes the theory from classical mental test theory, among others. If we consider the projection of the operating characteristic of a discrete item response on the criterion dimension, for example, then the resulting operating characteristic as a function of γ has to be *incidental*, for it has to be affected by the population distribution of θ.

We need to start from the conditional distribution of γ, given θ, therefore, which can be conceived of as being intrinsic in the relationship between the two variables, and independent of the population distribution of θ.

We assume that $\varsigma(\theta)$ takes on the same value only at a finite or an enumerable number of points of θ. Let $P^*_{k_g}(\varsigma)$ be the conditional probability assigned to the discrete response k_g, given ς. We can write

$$P^*_{k_g}(\varsigma) = \sum_{\varsigma(\theta)=\varsigma} P_{k_g}(\theta) \ . \tag{2.10}$$

III When $\varsigma(\theta)$ Is Strictly Increasing in θ : Simplest Case

[III.1] Amounts of Item and Test Information for a Fixed Value of ς

The simplest case is that $\varsigma(\theta)$ is strictly increasing in θ. In this case, $\varsigma(\theta)$ has a one-to-one correspondence with θ, and (2.10) becomes simplified into the form

$$P^*_{k_g}(\varsigma) = P^*_{k_g}[\varsigma(\theta)] = P_{k_g}(\theta) \ . \tag{3.1}$$

If, in addition, $\partial\theta/\partial\varsigma$ is finite throughout the entire range of θ, then we obtain

$$\frac{\partial}{\partial\varsigma} P^*_{k_g}(\varsigma) = [\frac{\partial}{\partial\theta} P_{k_g}(\theta)]\ \frac{\partial\theta}{\partial\varsigma} \ . \tag{3.2}$$

Let $I^*_{k_g}(\varsigma)$ be the item response information function defined as a function of ς. We can write

$$I^*_{k_g}(\varsigma) \ = \ -\frac{\partial^2}{\partial\varsigma^2} \log P^*_{k_g}(\varsigma) = -\frac{\partial}{\partial\varsigma}[\{\frac{\partial}{\partial\theta} \log P_{k_g}(\theta)\}\ \frac{\partial\theta}{\partial\varsigma}] \tag{3.3}$$

$$= \; I_{k_g}(\theta)\,(\frac{\partial\theta}{\partial\varsigma})^2 - [\frac{\partial}{\partial\theta}P_{k_g}(\theta)]\,[P_{k_g}(\theta)]^{-1}\,\frac{\partial^2\theta}{\partial\varsigma^2} \; .$$

Let $I_g^*(\varsigma)$ and $I^*(\varsigma)$ be the amounts of information given by a single item g and by the total test, respectively, for a fixed value of ς . Then we have from (2.3), (2.8) and (3.3)

$$I_g^*(\varsigma) = E[I_{k_g}^*(\varsigma) \mid \varsigma] = \sum_{k_g} I_{k_g}^*(\varsigma)\,P_{k_g}^*(\varsigma) = I_g(\theta)\,(\frac{\partial\theta}{\partial\varsigma})^2 \tag{3.4}$$

and

$$I^*(\varsigma) = \sum_{g=1}^{n} I_g^*(\varsigma) = I(\theta)\,(\frac{\partial\theta}{\partial\varsigma})^2 \; . \tag{3.5}$$

If we take the square roots of these two information functions defined for ς , then we obtain

$$[I_g^*(\varsigma)]^{1/2} = [I_g(\theta)]^{1/2}\,\frac{\partial\theta}{\partial\varsigma} \tag{3.6}$$

and

$$[I^*(\varsigma)]^{1/2} = [I(\theta)]^{1/2}\,\frac{\partial\theta}{\partial\varsigma} \; . \tag{3.7}$$

Since a certain constant nature exists for the square root of the item information function while the same is not true with the original item information function (cf. Samejima, 1979, 1982), $[I_g^*(\varsigma)]^{1/2}$ given by (3.6) instead of the original function given by (3.4) may be more useful in some occasions. This will be discussed later in this section, when the validity in selection *plus* classification is discussed.

[III.2] Validity in Selection

Suppose that we have a critical value, γ_0 , of the criterion variable, which is needed for succeeding in a specified job, and that we try to accept applicants whose values of the criterion variable are γ_0 or greater. If our primary purpose of testing is to make an accurate selection of applicants, then (3.6) and (3.7) for $\varsigma = \gamma_0$, or their squared values shown by (3.4) and (3.5), indicate item and test validities, respectively. In other words, if for some item formula (3.6) or (3.4) assumes high values at $\varsigma = \gamma_0$, then the standard error of estimation of ς around $\varsigma = \gamma_0$ becomes small and chances are slim that we make misclassifications of the applicants by accepting unqualified persons and rejecting qualified ones, and vice versa. The same logic applies to the total test by using formula (3.7) or (3.5) instead of (3.6) or (3.4).

It should be noted in (3.6) or in (3.7), that $[I_g^*(\tau)]^{1/2}$ or $[I^*(\tau)]^{1/2}$ consists of two factors, i.e., 1) the square root of the item information function $I_g(\theta)$ or that of the test information function $I(\theta)$ and 2) the partial derivative of ability θ with respect to ς at $\varsigma = \gamma_0$. These two factors in each formula are independent of each other, i.e., one belongs to the item or to the test and the other to the statistical relationship between θ and γ . We also notice that these two factors are in a supplementary relationship, i.e., even if one assumes a small value the other can supplement it in order to make the resulting product large. Thus while it is important to have a large amount of item information, or of test information, it is even more so to have large values of the derivative, $\partial\theta/\partial\varsigma$, in the vicinity of $\varsigma = \gamma_0$, for this will increase the amount of item information defined with respect to ς uniformly in that vicinity, and also that of test information, as is obvious from the right hand sides of (3.6) and

(3.7). In other words, it is desirable for the purpose of selection *for ς to increase slowly in θ in the vicinity of $\varsigma = \gamma_0$* .

Since, in general, the same ability θ has predictabilities for more than one kind of job performance, or of potential of achievement, *the performance function varies for different criterion variables.* Note that neither $[I_g(\theta)]^{1/2}$ nor $[I(\theta)]^{1/2}$ is changed even when the criterion variable is switched. Thus, for a fixed item or test whose amount of information is reasonably large around $\varsigma = \gamma_0$, the derivative $\partial\theta/\partial\varsigma$ in the vicinity of $\varsigma = \gamma_0$ determines the appropriateness of the use of the item or of the test for the purpose of selection with respect to a specific job, etc. If this derivative assumes a high value, then an item or a test which provides us with a medium amount of information may be acceptable for our purpose of selection, while we will need an item or a test whose amount of information is substantially larger if the derivative is low. Also for the same criterion variable γ the derivative $\partial\theta/\partial\varsigma$ varies for different values of γ_0 , so the appropriateness of an item or of a test depends upon our choice of γ_0 , too.

The above logic also applies for the formulae (3.4) and (3.5), i.e., for the case in which we choose the information functions, instead of their square roots, changing $\partial\theta/\partial\varsigma$ to its squared value.

It is obvious from (3.4) and (3.6) that we can choose either $I_g(\theta(\gamma_0))$ or $[I_g(\theta(\gamma_0))]^{1/2}$ for use in item selection, for their rank orders across different items are identical, and they equal the rank orders of $I_g^*(\gamma_0)$ as well as those of $[I_g^*(\gamma_0)]^{1/2}$.

[III.3] Validity in Selection Plus Classification

If we take another standpoint that our purpose of testing is not only to make a right selection of applicants but also to predict the degree of success in the job for each selected individual, then we will need to integrate $[I_g^*(\varsigma)]^{1/2}$ and $[I^*(\varsigma)]^{1/2}$, respectively, since we must estimate ς accurately not only around $\varsigma = \gamma_0$ but also for $\varsigma > \gamma_0$. If we choose $[I_g^*(\varsigma)]^{1/2}$ and $[I^*(\varsigma)]^{1/2}$ in preference to their squared values, we will obtain from (3.6) and (3.7)

(3.8) $$\int_{\Omega_\varsigma} [I_g^*(\varsigma)]^{1/2}\, d\varsigma = \int_{\Omega_\theta} [I_g(\theta)]^{1/2}\, d\theta$$

and

(3.9) $$\int_{\Omega_\varsigma} [I^*(\varsigma)]^{1/2}\, d\varsigma = \int_{\Omega_\theta} [I(\theta)]^{1/2}\, d\theta \ ,$$

where Ω_ς and Ω_θ indicate the domains of ς and θ for which $\varsigma(\theta) \geq \gamma_0$, respectively. In other words, when our purpose of testing is not only to make an accurate selection among the applicants but also to discriminate their ability accurately for future purposes among those who were accepted with respect to the criterion variable γ , we need to select items which assume high values of (3.8) instead of (3.6), or a test which provides us with a high value of (3.9) in place of (3.7).

Note that formulae (3.8) and (3.9) imply that we can obtain these two validity measures directly from the original item and test information functions, respectively, i.e., *without actually transforming θ to ς* , as long as we can identify the domain Ω_θ . This is true for any criterion variable γ .

Some examples illustrating the values of (3.8) are given in Figure 3-1 for hypothetical items. In the simplest case observed in this section and illustrated in Figures 2-1 and 3-1, these two domains, Ω_θ and Ω_ς , are provided by the two intervals, (θ_0 , ∞) and $(\gamma_0 , \bar{\gamma})$, where

(3.10) $$\theta_0 = \theta(\gamma_0)$$

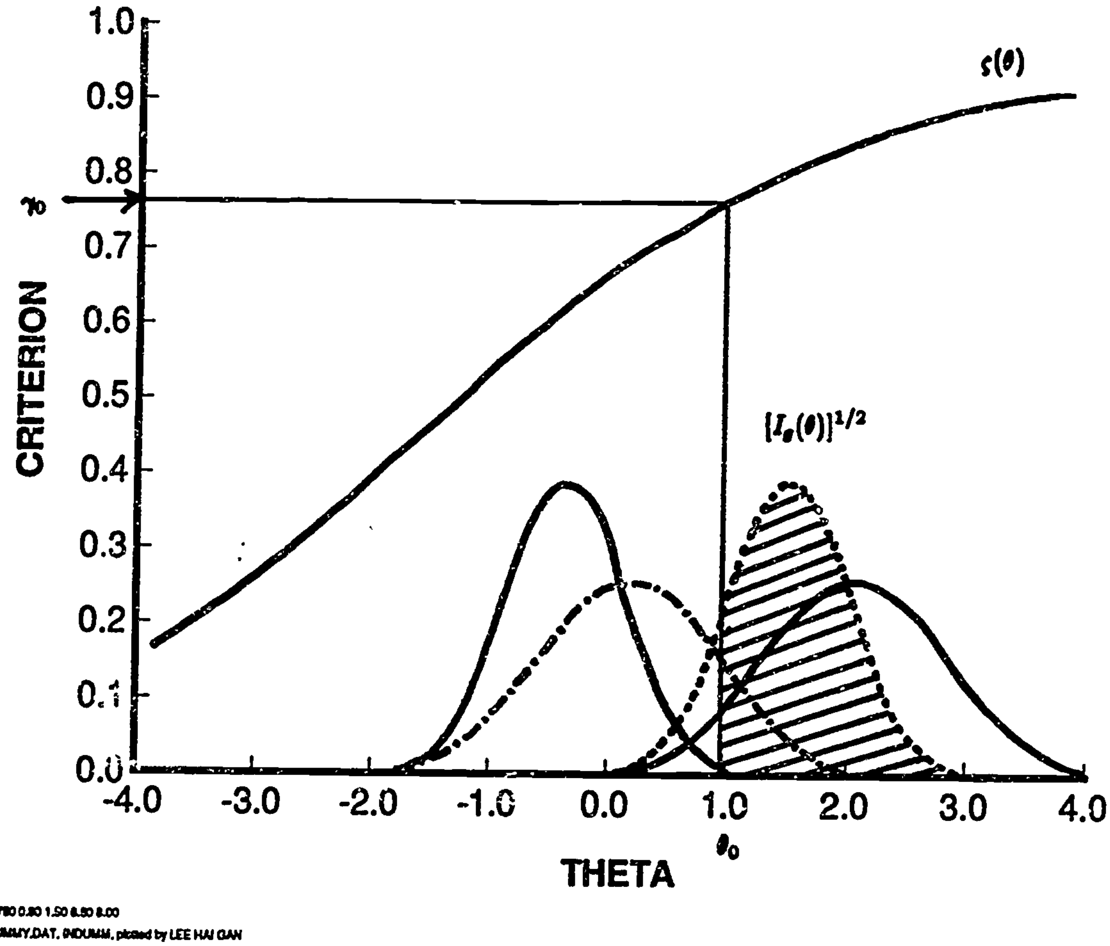

FIGURE 3-1

Some Examples of the Relationship between γ_0 and the Item Validity Measure Given by (3.8).

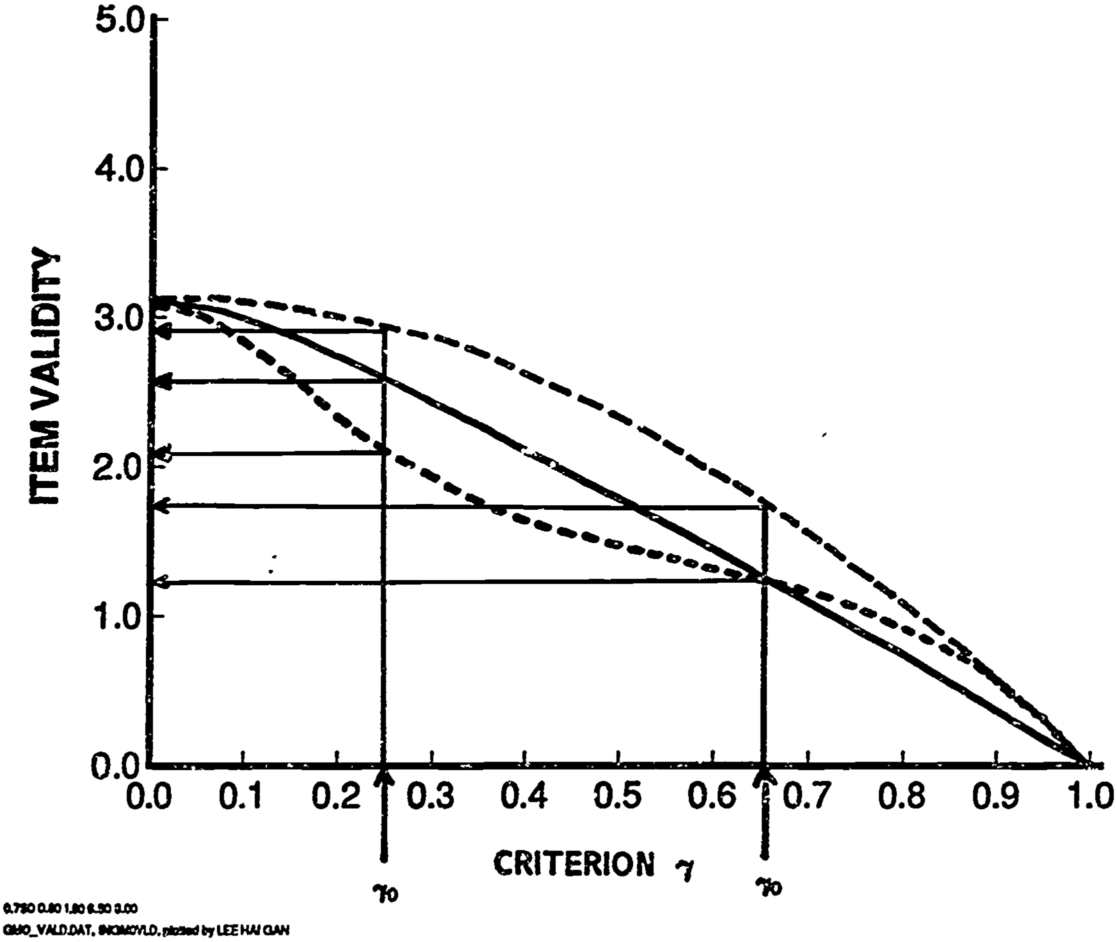

FIGURE 3-2

Relationship between γ_0 and Item Validity Indicated by (3.8) for Three Hypothetical Dichotomous Items Whose Operating Characteristics for the Correct Answer Are Strictly Increasing with Zero and Unity as Their Asymptotes.

and $\overline{\gamma}$ denotes the least upper bound of γ .

It should be noted that this pair of validity measures depends upon our choice of the critical value γ_0 . If this value is low, i.e., a specified job does not require high levels of competence with respect to the criterion variable γ , then these validity indices assume high values, and vice versa. It has been pointed out (Samejima, 1979, 1982) that there is a certain constancy in the amount of information provided by a single test item. To give some examples, if an item is dichotomously scored and has a strictly increasing operating characteristic for success with zero and unity as its two asymptotes, then the area under the curve for $[I_g(\theta)]^{1/2}$ equals π , *regardless of the mathematical form of the operating characteristic and its parameter values*; if it follows a three-parameter model with the lower asymptote, c_g (> 0) , then this area is less than π and strictly decreasing in, and solely dependent upon, c_g . We can see, therefore, that if our items belong to the first type then the functional relationship between γ_0 and the item validity measure given by (3.8) will be *monotone decreasing*, with π and zero as its two asymptotes, for each and every item. Figure 3-2 illustrates this relationship for three hypothetical items of this type. As we can see in this figure, the appropriateness of the items changes with γ_0 in an absolute sense, and also relatively to other items with γ_0 , and the rank orders of desirability among the items depend upon our choice of γ_0 .

We can see from (3.8) that this validity measure necessarily assumes a high value if an item is difficult, and the same applies to (3.9) for the total test. This implies that these validity measures alone cannot indicate the desirability of an item and of a test precisely *for a specific population of examinees.* In selecting items or a test, therefore, it is desirable to take the ability distribution of the examinees into account, if the information concerning the ability distribution of a *target* population is more or less available. In so doing we shall be able to avoid choosing items which are too difficult for the target population of examinees.

Let $f(\theta)$ denote the density function of the ability distribution for a specific population of examinees, and $f^*(\varsigma)$ be that of ς for the same population. Then we can write

$$f^*(\varsigma) = f(\theta) \frac{\partial\theta}{\partial\varsigma} \quad . \tag{3.11}$$

Adopting this as the weight function, from (3.6) and (3.7) we obtain as the validity indices *tailored* for a specific population of examinees

$$\int_{\Omega_\varsigma} [I_g^*(\varsigma)]^{1/2} f^*(\varsigma)\, d\varsigma = \int_{\Omega_\theta} [I_g(\theta)]^{1/2} f(\theta) \frac{\partial\theta}{\partial\varsigma}\, d\theta \tag{3.12}$$

and

$$\int_{\Omega_\varsigma} [I^*(\varsigma)]^{1/2} f^*(\varsigma)\, d\varsigma = \int_{\Omega_\theta} [I(\theta)]^{1/2} f(\theta) \frac{\partial\theta}{\partial\varsigma}\, d\theta \quad . \tag{3.13}$$

Thus by using (3.12) and (3.13) instead of (3.8) and (3.9) we shall be able to make appropriate item selection and test selection for a target population or sample, provided that the information concerning its ability distribution is more or less available. Note that, unlike (3.8) and (3.9), we need $\partial\theta/\partial\varsigma$ in evaluating these measures given by (3.12) and (3.13). Thus not only are these validity measures specific for the ability distribution of a target population, but also they are heavily dependent upon the functional formula of $\varsigma(\theta)$.

If we choose to use the area under the curve of the information function instead of that of its square root, we obtain from (3.4) and (3.5)

(3.14)
$$\int_{\Omega_\varsigma} I_g^*(\varsigma)\, d\varsigma = \int_{\Omega_\theta} I_g(\theta)\, \frac{\partial\theta}{\partial\varsigma}\, d\theta$$

and

(3.15)
$$\int_{\Omega_\varsigma} I^*(\varsigma)\, d\varsigma = \int_{\Omega_\theta} I(\theta)\, \frac{\partial\theta}{\partial\varsigma}\, d\theta \quad ,$$

respectively. We notice that in this case, unlike those of (3.8) and (3.9), the integrands of the right hand sides of (3.14) and (3.15) are no longer independent of the functional formula of $\varsigma(\theta)$. Also when information about the ability distribution of a target population of examinees is more or less available, the "tailored" item and test validity indices become

(3.16)
$$\int_{\Omega_\varsigma} I_g^*(\varsigma)\, f^*(\varsigma)\, d\varsigma = \int_{\Omega_\theta} I_g(\theta)\, f(\theta)\, (\frac{\partial\theta}{\partial\varsigma})^2\, d\theta$$

and

(3.17)
$$\int_{\Omega_\varsigma} I^*(\varsigma)\, f^*(\varsigma)\, d\varsigma = \int_{\Omega_\theta} I(\theta)\, f(\theta)\, (\frac{\partial\theta}{\partial\varsigma})^2\, d\theta \quad ,$$

respectively, if we choose to use the infomation functions instead of their square roots.

Note that, unlike the validity measures for "selection" purposes, in the present situation the rank orders of validity across different items, or different tests, depend upon the choice of the validity index. Thus a question is: which of the formulae, (3.8) or (3.14), and (3.9) or (3.15), are better as the item and the test validity indices for "selection plus classification" purposes? A similar question is also addressed with respect to (3.12) and (3.16), and to (3.13) and (3.17). These are tough questions to answer. While the choice of the square root of the item information function has an advantage of a certain constancy which has been observed earlier in this subsection, the use of the item information has a benefit of additivity, i.e., by virtue of (2.8) the sum total of (3.14) over all the item g 's equals (3.15), and the same relationship holds between (3.16) and (3.17). The answers to these questions are yet to be searched.

[III.4] Validity in Classification

When our purpose of testing is strictly the classification of individuals, as in assigning those people to different training programs, in guidance, etc., (3.8) and (3.9), or (3.14) and (3.15), also serve as the validity measures of an item and of a test, respectively. In this case, we must set $\gamma_0 = \underline{\gamma}$ in defining the domains, Ω_ς and Ω_θ, where $\underline{\gamma}$ is the greatest lower bound of γ. Thus the two domains, Ω_ς and Ω_θ, in these formulae become those of ς and θ for which $\underline{\gamma} \le \varsigma(\theta) \le \overline{\gamma}$. It is obvious that these formulae provide us with the item and the test validity measures, respectively, for the same reason explained in [III.3].

The same logic applies for the "tailored" validity measures provided by (3.12) and (3.13), and by (3.16) and (3.17), when the information concerning the abi[illegible]y distribution of a target population is more or less available.

[III.5] Computerized Adaptive Testing

The item information function, $I_g(\theta)$, has been used in the computerized adaptive testing in selecting an optimal item to *tailor* a sequential subtest of items for an individual examinee out of the prearranged itempool. A procedure may be to let the computer choose an item having the highest value of $I_g(\theta)$ at the current estimated value of θ for the individual examinee, which is based upon his responses to the items that have already been presented to him in sequence, out of the set of remaining items in the itempool.

We notice from (3.4) or (3.6) that this procedure is justified from the standpoint of criterion-oriented validity, for the item which provides us with the greatest item information $I_g(\theta)$ among all the available items in the itempool also gives the greatest values of $I_g^*(\varsigma)$ and its square root, at any fixed value of θ .

Amount of test information can be used effectively in the stopping rule of the computerized adaptive testing. A procedure may be to terminate the presentation of a new item out of the itempool to the individual examinee when $I(\theta)$ has reached an a priori set amount at the current value of his estimated θ .

When we have a specific criterion variable γ in mind, it is justified to use an a priori set value of $I^*(\varsigma)$ instead of $I(\theta)$. In so doing we can obtain the value of $I(\theta)$ corresponding to the a priori set value of $I^*(\varsigma)$ for each θ , through the formula

(3.18)
$$I(\theta) = I^*(\varsigma)\ (\frac{\partial \varsigma}{\partial \theta})^2 \ ,$$

which is obtained from (3.7). Thus it is easy to have the computer to handle this situation, provided that we know the functional formula for $\varsigma(\theta)$.

IV Test Validity Measures Obtained from More Accurate Minimum Variance Bounds

When $\{\partial\varsigma/\partial\theta\} = 0$ at some value of θ , as is illustrated by a dashed line in Figure 2-2, $\partial\theta/\partial\varsigma$ becomes positive infinity, and so does the item validity measure given by (3.6). This fact provides us with some doubt, for, while we can see that at such a point of ς item validity is high, we must wonder if positive infinity is an adequate measure. It is also obvious from (2.8) that the same will happen to the total test if it includes at least one such item. Our question is: *should we search for more meaningful functions than the item and test information functions?* This topic will be discussed in this section.

Necessity of the search for a more accurate measure than the test information function becomes more urgent when the performance function, $\varsigma(\theta)$, is not strictly increasing in θ , but is, say, only piecewise monotone in θ with finite $\partial\theta/\partial\varsigma$ and differentiable with respect to θ , as is illustrated in Figure 4-1. The illustrated performance function is still simple enough, but indicates the trend that after a certain point of ability the performance level in a specified job decreases. This can happen when the job does not provide enough challenge for persons of very high ability levels.

Since $I^*(\varsigma)$ serves as the reciprocal of the conditional variance of the maximum likelihood estimate of ς only asymptotically and there exist more accurate minimum variance bounds for any (asymptotically) unbiased estimator (cf. Kendall and Stuart, 1961), we can search for more accurate test validity measures than the one given by (3.7) by using the reciprocal of the square roots of such minimum variance bounds. Details of this topic will be discussed in a separate paper. Here its brief summary related to validity measures will be given.

Let $J_{rs}(\theta)$ be defined as

(4.1)
$$J_{rs}(\theta) = E[\frac{L_V^{(r)}(\theta)}{L_V(\theta)}\ \frac{L_V^{(S)}(\theta)}{L_V(\theta)}\ |\ \theta] \qquad r, s = 1, 2, ..., k$$

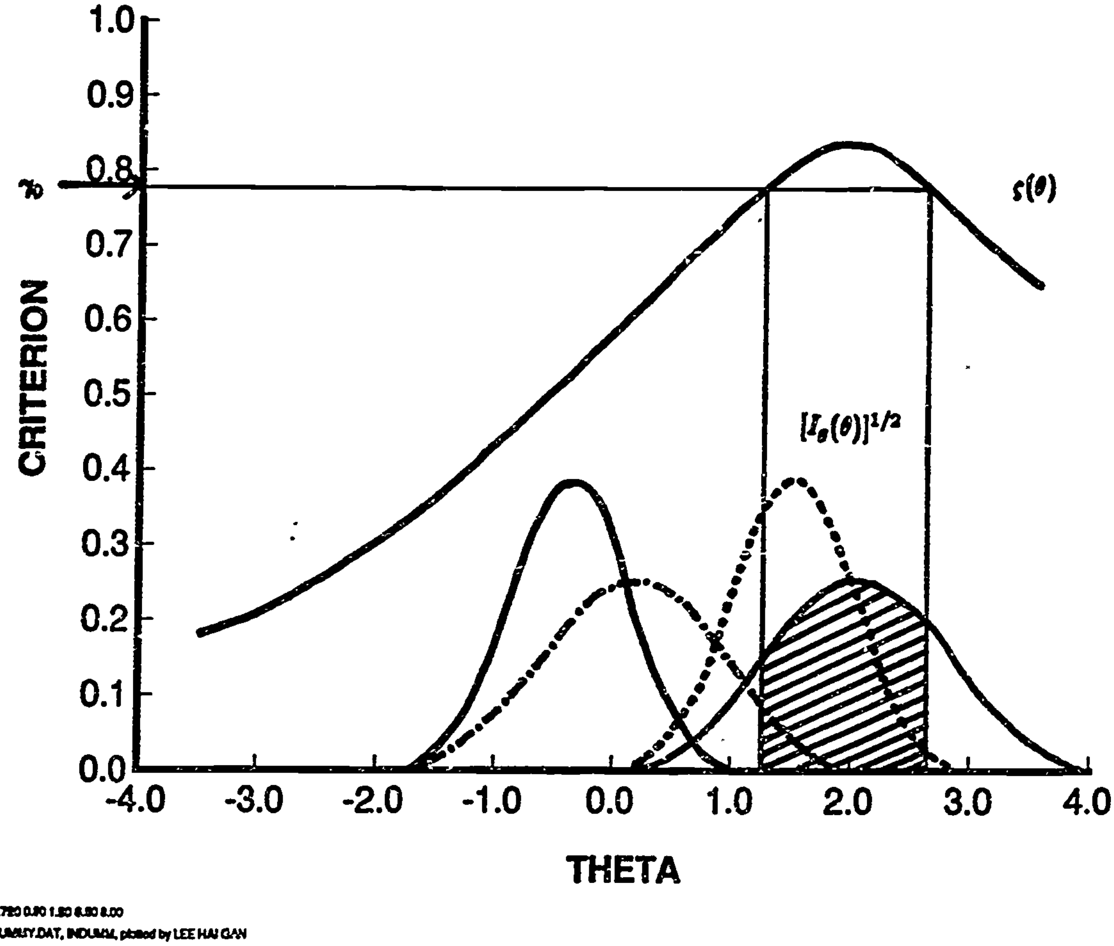

FIGURE 4-1

Example of the Performance Function $\varsigma(\theta)$ Which Is Piecewise Monotone in θ .

18

where

$$(4.2) \qquad L_V^{(r)}(\theta) = \frac{\partial^r}{\partial\theta^r} L_V(\theta) = \frac{\partial^r}{\partial\theta^r} P_V(\theta) \ .$$

Let $J(\theta)$ denote the (k x k) matrix of the element $J_{rs}(\theta)$, and $J_{rs}^{-1}(\theta)$ be the corresponding element of its inverse matrix, $J^{-1}(\theta)$. Note that when $k=1$ we can rewrite (4.1) into the form

$$(4.3) \qquad \begin{aligned} J_{kk}(\theta) &= J_{11}(\theta) = E[\{\frac{\partial}{\partial\theta}\log L_V(\theta)\}^2 \mid \theta] \\ &= -E[\frac{\partial^2}{\partial\theta^2}\log P_V(\theta) \mid \theta] \ , \end{aligned}$$

and from this, (2.7) and (2.8) we can see that $J(\theta)$ is a (1 x 1) matrix whose element is the test information function, $I(\theta)$, itself. A set of improved minimum variance bounds is given by

$$(4.4) \qquad \sum_{r=1}^{k}\sum_{s=1}^{k} \varsigma^{(s)}(\theta)\ J_{rs}^{-1}(\theta)\ \varsigma^{(r)}(\theta)$$

(cf. Kendall and Stuart, 1961), where $\varsigma^{(s)}(\theta)$ denotes the s-th partial derivative of $\varsigma(\theta)$ with respect to θ . We obtain, therefore, for a set of new test validity measures

$$(4.5) \qquad [\sum_{r=1}^{k}\sum_{s=1}^{k} \gamma_0^{(s)}\ J_{rs}^{-1}(\theta(\gamma_0))\ \gamma_0^{(r)}]^{-1/2} \ ,$$

where $\gamma_0^{(s)}$ indicates the s-th partial derivative of ς with respect to θ at $\varsigma = \gamma_0$.

The use of this new test validity measure will ameliorate the problems caused by $\{\partial\varsigma/\partial\theta\} = 0$, if we choose an appropriate k . The resulting algorithm will become much more complicated, however, and we must expect a substantially larger amount of CPU time for computing these measures when k is greater than unity. Note that (4.5) equals (3.7) when $k = 1$.

V Multidimensional Latent Space

When our latent space is multidimensional, a generalization of the idea given in Section 4 for the unidimensional latent space can be made straightforwardly. We can write

$$(5.1) \qquad \theta = \{\ \theta_u\ \}' \qquad u = 1, 2, ..., \eta \ ,$$

and the performance function $\varsigma(\theta)$ becomes a function of η independent variables. A minimum variance bound is given by

$$(5.2) \qquad \sum_{u=1}^{\eta}\sum_{v=1}^{\eta} \frac{\partial\varsigma(\theta)}{\partial\theta_u}\ \frac{\partial\varsigma(\theta)}{\partial\theta_v}\ I_{uv}^{-1}(\theta) \ ,$$

where $I_{uv}^{-1}(\theta)$ is the (u,v)-element of the inverse matrix of the $(\eta\ \text{x}\ \eta)$ symmetric matrix, whose element is given by

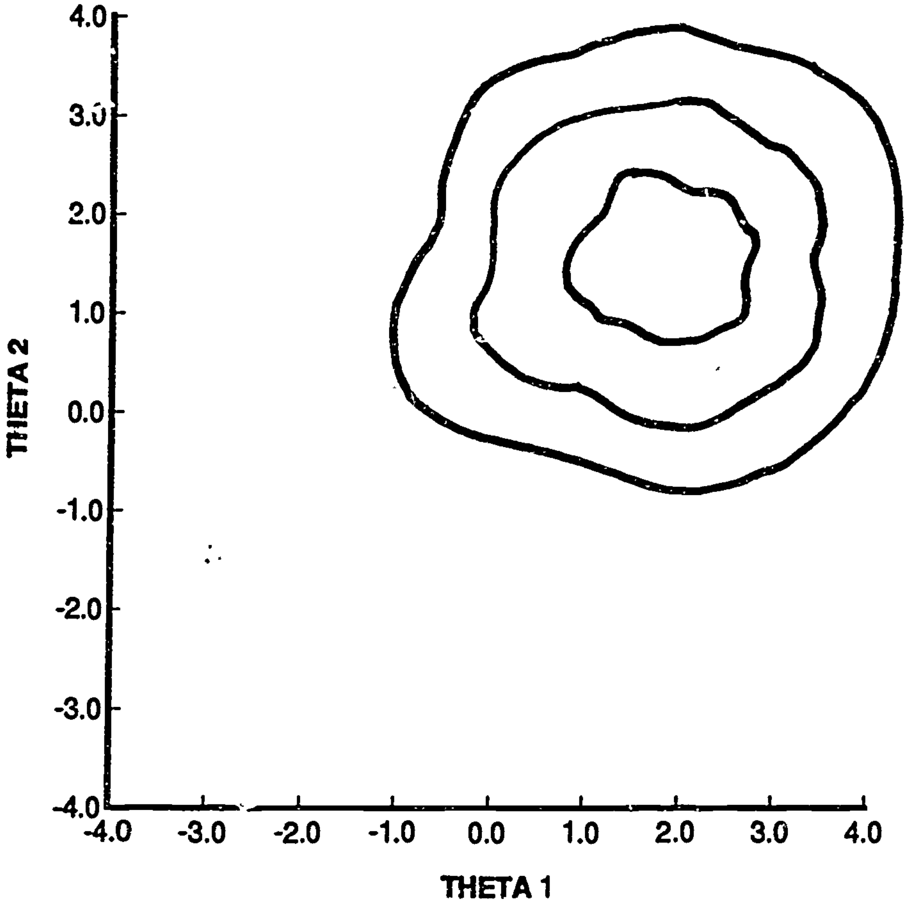

FIGURE 5-1

Area Ω_θ for Different γ_0's in Two-Dimensional Latent Space for a Hypothesised Test.

(5.3) $$I_{uv}(\theta) = E[\frac{1}{L}\frac{\partial L}{\partial \theta_u}\frac{\partial L}{\partial \theta_v} \mid \theta]$$

with L abbreviating $L_V(\theta)$, or $P_V(\theta)$. The reciprocal of the square root of (5.3) will provide us with the counterpart of (3.7) for the multidimensional latent space. For $\eta = 2$, the area Ω_θ may look like one of the contours illustrated in Figure 5-1, depending upon our choice of γ_0 , taking the axis for γ vertical to the plane defined by θ_1 and θ_2 .

In a more complex situation where both ability and the criterion variables are multidimensional, we must consider the projection of the item information function on the criterion subspace from the ability subspace, in order to have the item validity function for each item, and then the test validity function. It is anticipated that we must deal with a higher mathematical complexity in such a case. The situation will substantially be simplified, however, if the total set of items consists of several subsets of items, each of which measures, exclusively, a single ability dimension and a single criterion dimension.

VI Discussion and Conclusions

In contrast to the progressive desolution of the reliability coefficient in classical mental test theory and the replacement by the test information function in latent trait models, the issue of test validity has been more or less neglected in modern mental test theory. The present paper proposes some considerations about the validity of a test and of a single item. Effort has been focused upon searching for measures which are population-free, and which will provide us with local and abundant information just as the information functions do in comparison with the test reliability coefficient in classical mental test theory. In so doing, validity indices for different purposes of testing and also those which are tailored for a specific population of examinees are considered.

The above considerations for the item and test validities may be just part of many possible approaches. We may still have a long way to go before we discover the most useful measures of the item and test validities. The aim of the present paper is rather to provide stimulation so that researchers will pursue this topic further, taking different approaches.

References

[1] Kendall, M. G. and Stuart, A. *The advanced theory of statistics. Vol. 2.* New York: Hafner, 1961.

[2] Samejima, F. A use of the information function in tailored testing. *Applied Psychological Measurement*, 1, 1977, 233-247.

[3] Samejima, F. Constant information model: A new promising item characteristic function. *ONR/RR-79-1*, 1979.

[4] Samejima, F. Information loss caused by noise in models for dichotomous items. *ONR/RR-82-1*, 1982.

[5] Samejima, F. Predictions of reliability coefficients of a test using the test information function. *ONR/RR-90-3*, 1990.

ONRR9003.TEX
June 14, 1990

Distribution List

Dr. Terry Ackerman
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

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

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

Dr. Ronald Armstrong
Rutgers University
Graduate School of Management
Newark, NJ 07102

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

Dr. Laura L. Barnes
College of Education
University of Toledo
2801 W. Bancroft Street
Toledo, OH 43606

Dr. William M. Bart
University of Minnesota
Dept. of Educ. Psychology
330 Burton Hall
178 Pillsbury Dr., S.E.
Minneapolis, MN 55455

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

5/1/90

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

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

Dr. Bruce Bloxom
Defense Manpower Data Center
99 Pacific St.
Suite 155A
Monterey, CA 93943-3231

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

Dr. Robert Breaux
Code 281
Naval Training Systems Center
Orlando, FL 32826-3224

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

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

Dr. John M. Carroll
IBM Watson Research Center
User Interface Institute
P.O. Box 704
Yorktown Heights, NY 10598

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

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

Mr. Hua Hua Chung
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

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

Director, Manpower Program
Center for Naval Analyses
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

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

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

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

Ms. Carolyn R. Crone
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Timothy Davey
American College Testing Program
P.O. Box 168
Iowa City, IA 52243

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

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

Dr. Lou DiBello
CERL
University of Illinois
103 South Mathews Avenue
Urbana, IL 61801

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
(12 Copies)

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

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

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

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

ERIC Facility-Acquisitions
2440 Research Blvd, Suite 550
Rockville, MD 20850-3238

Dr. Benjamin A. Fairbank
Operational Technologies Corp.
5825 Callaghan, Suite 225
San Antonio, TX 78228

Dr. Marshall J. Farr, Consultant
Cognitive & Instructional Sciences
2520 North Vernon Street
Arlington, VA 22207

Dr [illegible] Federico
Code 51
NPRDC
San Diego, CA 92152-6800

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

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

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

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

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

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

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

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

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

Dr. Drew Gitomer
Educational Testing Service
Princeton, NJ 08541

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

Dr. Sherrie Gott
AFHRL/MOMJ
Brooks AFB, TX 78235-5601

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

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

Prof. Edward Haertel
School of Education
Stanford University
Stanford, CA 94305

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

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

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

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

Dr. Thomas M. Hirsch
ACT
P. O. Box 168
Iowa City, IA 52243

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

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

Ms. Julia S. Hough
Cambridge University Press
40 West 20th Street
New York, NY 10011

Dr. William Howell
Chief Scientist
AFHRL/CA
Brooks AFB, TX 78235-5601

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

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

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

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

Dr. Kumar Joag-dev
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright Street
Champaign, IL 61820

Dr. Douglas H. Jones
1280 Woodfern Court
Toms River, NJ 08753

5/1/90

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

Dr. Michael Kaplan
Office of Basic Research
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

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

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

Dr. Jwa-keun Kim
Department of Psychology
Middle Tennessee State University
P.O. Box 522
Murfreesboro, TN 37132

Mr. Soon-Hoon Kim
Computer-based Education Research Laboratory
University of Illinois
Urbana, IL 61801

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

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

Dr. Richard J. Koubek
Department of Biomedical & Human Factors
139 Engineering & Math Bldg.
Wright State University
Dayton, OH 45435

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

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

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

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

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

Mr. Rodney Lim
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

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

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

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

Dr. Richard Luecht
ACT
P. O. Box 168
Iowa City, IA 52243

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

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

Dr. Clessen J. Martin
Office of Chief of Naval Operations (OP 13 F)
Navy Annex, Room 2832
Washington, DC 20350

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

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

Mr. Christopher McCusker
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert McKinley
Educational Testing Service
Princeton, NJ 08541

Mr. Alan Mead
c/o Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Timothy Miller
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

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

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

Headquarters Marine Corps
Code MPI-20
Washington, DC 20380

Dr. Ratna Nandakumar
Educational Studies
Willard Hall, Room 213E
University of Delaware
Newark, DE 19716

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

Librarian
Naval Center for Applied Research in Artificial Intelligence
Naval Research Laboratory
Code 5510
Washington, DC 20375-5000

5/1/90

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

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

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

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

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

Dr. Peter J. Pashley
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

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

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

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

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

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

Mr. Steve Reiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

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

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

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

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

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

Dr. Mary Schratz
905 Orchid Way
Carlsbad, CA 92009

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

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

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

Dr. Randall Shumaker
Naval Research Laboratory
Code 5510
4555 Overlook Avenue, S.W.
Washington, DC 20375-5000

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

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

Dr. 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
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

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

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

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

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

Dr. Kikumi Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

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. Thomas J. Thomas
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

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

5/1/90

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

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

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

Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

Dr. Michael T. Waller
University of Wisconsin-Milwaukee
Educational Psychology Department
Box 413
Milwaukee, WI 53201

Dr. Ming-Mei Wang
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Thomas A. Warm
FAA Academy AAC934D
P.O. Box 25082
Oklahoma City, OK 73125

Dr. Brian Waters
HumRRO
1100 S. Washington
Alexandria, VA 22314

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

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

5/1/90

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

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

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

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

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

Dr. Hilda Wing
Federal Aviation Administration
800 Independence Ave, SW
Washington, DC 20591

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

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

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

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

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

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

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