AD-A113 578 NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN D--ETC F/G 5/9
ON MODELS AND METHODS FOR PERFORMANCE MEASUREMENT.(U)
MAR 82 W J MOONAN
UNCLASSIFIED NPRDC-TN-82-11 NL

END
DATE
FILMED
5-82
DTIC

12

NPRDC TN 82-11

MARCH 1982

ON MODELS AND METHODS FOR PERFORMANCE MEASUREMENT

AD A113578

NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN DIEGO, CALIFORNIA

NAVY PERSONNEL RESEARCH
AND
DEVELOPMENT CENTER
San Diego, California 92152

82 04 19 009

NPRDC Technical Note 82-11 March 1982

ON MODELS AND METHODS FOR PERFORMANCE MEASUREMENT

William J. Moonan

Reviewed by
Jules Borack

Released by
James F. Kelly, Jr.
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER NPRDC TN 82-11	2. GOVT ACCESSION NO. AD-A113 578	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON MODELS AND METHODS FOR PERFORMANCE MEASUREMENT		5. TYPE OF REPORT & PERIOD COVERED Final Report Aug 1979-Sep 1980
		6. PERFORMING ORG. REPORT NUMBER 11-80-12
7. AUTHOR(s) W. J. Moonan		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS IED ZF66-512-001-03.06
11. CONTROLLING OFFICE NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		12. REPORT DATE March 1982
		13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Measurement models
Performance measurement
Logistic
Job performance
Research model

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

Models were specified and methods defined that characterize performance measurement as a process and as a function of (1) the performer's capability, (2) variables that indicate how well job operations are performed, and (3) the difficulties of those operations. The report shows methods of mathematical measurement model development that relate (1), (2), and (3) above. A performance quality model was defined. Illustrations of two performance models were provided through examples.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

FOREWORD

This research was conducted under the NAVPERSRANDCEN independent exploratory development (IED) program in analytical personnel research techniques. The objective of the work described in this report is to examine the mathematical basis for performance measurement and the ways in which it affects applications. Particular emphasis is placed on the mathematical relationship among performer capability, task variables, and task difficulty as it relates to models of performance measurement.

Results of the research are intended for use by researchers in the evaluation of individual performance.

JAMES F. KELLY, JR.
Commanding Officer

JAMES J. REGAN
Technical Director

SUMMARY

Problem

Performance measurement is antecedent and fundamental to the conduct of personnel research and personnel operations in military systems. It has a significant role to play in regard to personnel acquisition, training, duty, assignment, classification, and separation. In view of the importance of performance measurement to the military services, new methods of measuring performance that may be conveniently and economically utilized in operational situations must be developed.

Objectives

The objectives of this research were to (1) delineate the essential factors of performance measurement, (2) distinguish methods for developing performance measurement models, (3) indicate how mathematical models for performance measurement are developed from those methods, and (4) illustrate how these models can be applied in operational and practical situations to obtain performance measurements.

Approach

To develop a method to measure a performer's capability, the factors affecting the measurement must be identified. The critical element in the process is the comparison between the performer's capability and the difficulty of the task to be performed. This comparison and the observable information obtainable when the performer actually performs the operations of a job are sufficient for defining mathematical measurement models of a person's capability for several types or mixtures of performance variables.

Findings

Three methods of model development are identified and two models are specified. Model I uses observable information expressed qualitatively and Model II uses observable information expressed quantitatively. This latter model accommodates operations that are associated with variables having different distributional forms, including a mixture of qualitative and quantitative operational variables. Examples are provided to show how each model can be applied in practice.

Conclusions

1. Model I is a performance model that can be efficiently used to measure performance by means of test items. It is simple to understand and to use, and has the property of specific objectivity.

2. Model II is a performance model that can be effectively used to measure performance of persons in jobs requiring a variety of operations, each possessing its own form of observational variable distribution.

3. Job performance measurement models are conveniently developed as functions of the comparison between the performer's capability and the difficulties of the operations to be performed on the job.

Recommendations

1. Model I should be used to assess abilities of persons who are to be evaluated by tests in which items are scored with dichotomous responses.

2. Model II should be used to assess capabilities of persons to perform jobs whose operations are evaluated by means of variables with different distributional forms.

CONTENTS

INTRODUCTION

Problem

Performance measurement (PM) is viewed as fundamental to the conduct of personnel research and operations in all military systems. Perhaps no organized group of scientists appreciates this more than do personnel research workers. In regard to personnel, PM has a significant role to perform in the following areas: (1) acquisition, (2) training, (3) duty, (4) assignment, (5) classification, and (6) separation.

Objectives

The objectives of this research were to (1) delineate the essential factors of performance measurement, (2) distinguish methods for developing performance measurement models, (3) indicate how mathematical models for performance measurement are developed from those methods, and (4) illustrate how these models can be applied in operational and practical situations to obtain peformance measurements.

Background

Consider the logistic function of some variable t(ni) to be defined as

$$f\ t(ni) = \frac{1}{1 + \exp\left[-t(ni)\right]} = \frac{\exp\left[t(ni)\right]}{1 + \exp\left[t(ni)\right]} = \exp\left[t(ni)\right]/D(ni) \qquad (1)$$

From the table below, it can be noted that equation (1) "behaves" like the probability distribution function (also referred to as the cumulative distribution function (cdf) of t(ni)). This function expresses the probability that the variable t may take on a value less than or equal to t(ni). This may be expressed as

$$f\left[t(ni)\right] = \Pr\left[t \le t(ni)\right]. \qquad (2)$$

Equation (1) may serve as a distribution function provided it is nondecreasing with increasing t(ni) and satisfies $f(-\infty) = 0$ and $f(+\infty) = 1$. There are some other technical matters to be considered in order to be mathematically precise (continuity and integrability). However, for the present purposes, these are ignored. Numerical evaluation of (1) shows the following:

t(ni)	exp [t(ni)]	D(ni)	Pr [t(ni)]
$-\infty$	.00000	1.00000	.00000
-10	.00004	1.00004	.00004
-5	.00674	1.00674	.00669
-2	.13534	1.13534	.11920
-1	.36788	1.36788	.26894
0	1.00000	2.00000	.50000
1	2.71828	3.71828	.73106
2	7.38906	8.38906	.88080
5	148.41316	149.41316	.99331
10	22026.456	22027.456	.99995
$+\infty$	∞	∞	1.00000

Equation (1) can be shown to have other desirable properties. For example, if (2) is abbreviated as:

$$P(ni) = Pr\left[t<t(ni)\right] = \exp\left[t(ni)\right]/D(ni)\text{, and} \quad (3)$$

$$Q(ni) = 1 - P(ni) = 1/D(ni), \quad (4)$$

then,

$$\frac{P(ni)}{Q(ni)} = \exp\left[t(ni)\right]\text{, and} \quad (5)$$

the natural logarithm of (5) is

$$L(ni) = \ln\left[\frac{P(ni)}{Q(ni)}\right] = t(ni). \quad (6)$$

Consider also two different values of i, say i and j, $i \neq j$:

$$P(ni)\cdot Q(nj) = \exp\left[t(ni)\right]/D(ni)\cdot D(nj)\text{ , and} \quad (7)$$

$$P(nj)\cdot Q(ni) = \exp\left[t(nj)\right]/D(ni)\cdot D(nj). \quad (8)$$

Then,

$$R(nij) = \frac{P(ni)\cdot Q(nj)}{P(nj)\cdot Q(ni)} = \exp\left[t(ni) - t(nj)\right] . \quad (9)$$

Finally, the natural logarithm of R(nij) is

$$L(nij) = t(ni) - t(nj). \quad (10)$$

For particular definitions of t(ni), these and other properties of (1) are very useful.

PERFORMANCE MEASUREMENT

Measurement Process

Earlier, Panos (1969) remarked that the process of taking measurements is a process of making comparisons. This concept has been augmented by the author with a feature that contributes to understanding of the topic of "measurement." More explicit information about this feature will be reported later in the report.

The technical discussion of PM begins with two short and innocuous statements related to PM:

1. Statement A. The evaluator's goal in PM is to measure, on a quantitative scale, the property of a performer, called capability, that performers exhibit while interacting with contexts sampled from a situation of interest to the evaluator.

2. Statement B. A performer is said to have capability to the extent that the interactions with the contexts exhibit information by which the level of capability may be ascertained.

From these statements, the essential factors from which models of PM may be constructed will be identified.

Factors of PM

The Situation

Statement A implies that an evaluator's purpose is to measure a performer's capability with respect to some situation of interest to the evaluator. A "situation" is a set of conditions, activities, or circumstances (contexts) that collectively characterize an important or relevant state of affairs. To be more definite, important situations are often "jobs" and the contexts of these are the set of operations that must be undertaken to perform each job. Sometimes a training program is a situation and contexts are particular types of learning.

Performer Capability

Both statements refer to a property of a performer, called his capability. We assume that this property controls how well persons perform the jobs they undertake. A liberal view of this concept would be that it might also be a function of the performer's motivation (or other psychological factors) that might significantly influence performance. These other factors are neither separately identified nor recognized in the models developed in this report. It is assumed that the capability sufficiently represents the concept.

The capability of person (n) is such an important component in PM that it is a parameter of PM models and is given the symbol $\alpha(n)$. Its measurement is the raison d'être for the PM process.

Operation Difficulty

The statements also imply that the operations performed while doing a job have some difficulty, or else persons would not require a capability in order to perform them. The "difficulty" of operation (i) is also a parameter and is given the symbol $\delta(i)$.

Comparison

In the Measurement Process section, the concept of a comparison was mentioned. We define the relevant comparison for the PM process as

$$t(ni) = \alpha(n) - \delta(i). \tag{11}$$

This is the feature that contributes to an understanding of the PM process. If the difficulty of an operation is high, the concept presented here is that $\alpha(n)$ must be even greater if the performer is likely to perform the operation effectively. Performances associated with negative values of t(ni) have, by (2), a relatively small probability of good performance on the given operation.

The Observables

Statement B indicates that a person's performance with respect to an operation results in some information, usually behavioral evidence, which indicates, in some sense, the capability of the performer with respect to the difficulty of the operation. When appropriate, the characterization of this information is symbolized for person (n) and operation (i) as X(ni) or X(i).

It is the contention of this report that the three factors, $\alpha(n)$, $\delta(i)$, and X(ni), are the essential factors for developing models for PM. The factor $\alpha(n)$ specifies a latent (unobservable) trait of a performer, $\delta(i)$ characterizes a latent trait of the operation performed, and X(ni) provides an observable link that may mathematically relate the first two factors and, hence, provide a means by which each may be estimated and interpreted. Any capability afforded by PM to estimate $\alpha(n)$ quantitatively is extremely significant for personnel research as it immediately provides measurement for which the powerful tools of quantitative statistical methods may be employed for solving problems associated with the aspects of personnnel research and operations mentioned in the Introduction.

The observable performance information may be characterized in many different ways. Five of these ways are listed below.

1. Performance Quality. The outcome of a performance may be considered from the point of view of its quality. The quality may be thought of as being established on the basis of judgment or by more objective evidence. In any event, the quality of performance may be indicated by classification. The performance may be classified in one of K categories of quality. One of the PM models reported here uses exactly this type of observation.

2. Performance Quantity. Similarly, the fortunate evaluator may acquire quantitative information. Such may also be utilized in a PM model. In this case, the measurement problem is primarily a statistical problem of ascertaining the statistical distribution of the X(ni). Likely candidates are the normal, beta, and gamma distributions as variations in their parameters provide the great variety for forms observed in practice.

3. Performance Counts. Still another type of variable that might represent performance information are variables resulting from counting operations. Such counts may either represent performance errors or attainments or both. Such counting data may have Poisson distributions.

4. Performance Time. Performance time is an obvious choice for characterizing performance information. Its distribution is difficult to specify in general, but it seems likely that one of the distributions mentioned previously in these descriptions may be appropriate.

5. Mixed Performance Variables. Some situations may allow each operation's performance to be represented by performance variables of the same type, but, in general, this seems unlikely. Usually, a "job" requires that several operations be performed and each may be characterized by a different type of performance variable. The method described in this report of measuring $\alpha(n)$ must and does allow for this circumstance to occur. This facility is permitted because each performance variable is transformed to a common metric--its cumulative probability that the mathematical model of the next section relates to the PM parameters $\alpha(n)$ and $\delta(i)$ independently of the type of performance variable.

PM Models

Three categories of methods for developing measurement models are presented.

Method I

This form was originally defined and used by Rasch (1960) to develop a measurement model for measuring capability by observing dichotomous responses made by performers of operations (achievement test items). The feature of note here is that the Method I type of PM model incorporates the comparison (11) and the observable $X(ni)$ into the same expression. This, in turn, endows the model with a property called "objectivity," meaning that the $\alpha(n)$ estimates are independent of which operations are used and that $\delta(i)$ estimates are independent of which performers are involved (see Wright, 1967 and Choppin, 1968).

Method II

The form of this type of PM model is introduced in this report. Its distinctive aspect is that it is recognized that the function (1) of the comparison (11) is a cumulative probability and this probability is set equal to the percentile rank (another name for cumulative probability) of an observed performance variable. The unique idea is that, instead of assuming that a function of observed performance variables is equal to capability, the model assumes that that observable function equals a function of the comparison (11). In effect, the model says that the performance observed depends upon the difference between the capability of the performer and the difficulty of the operation being performed and not only upon the performer's capability.

Method III

This form, introduced by Wright and Mead (1978), substitutes (1) for the probability in the distribution functions of the performance variable that is observed. The authors do this for a variable with a binomial distribution in their report.

RESULTS AND DISCUSSION

Comparison, Capability, and Difficulty Estimation

For purposes of continuity, an application of Model II is provided before addressing the application of Model I. Next, we will show how to estimate (11) for a Method II model and the two parameters comprising (11). First, we have by (6) that

$$L(ni) = \alpha(n) - \delta(i). \tag{12}$$

Equation (12) implies that $\alpha(n)$ does not depend upon (i). This, in turn, implies that Model II assumes the existence of a single job performance capability and not a multidimensional set of abilities--one for each job operation. Those separate abilities may be estimated by applying Model II to each job operation separately. In general, if a job consists of I operations, (12) can be summed over these operations yielding from (12):

$$\sum_{i=1}^{I} L(ni) = I\alpha(n) - \sum_{i=1}^{I} \delta(i). \tag{13}$$

Therefore, we can estimate $\alpha(n)$ for a performer from

$$\hat{\alpha}(n) = \sum_{i=1}^{I} L(ni) + \hat{\delta}(i) \;/I \tag{14}$$

where the person performs I operations of known difficulty and L(ni) is estimated from (6). To use (14), we need to show how to estimate the $\delta(i)$. Equation (14) simplifies considerably if we can assume (19) and that the "job ability" is the same for all job operations.

Consider, for simplicity, that a "job" is comprised of I=3 operations, the performance variables being, perhaps, of I=3 different types. We then observe X(ni) on each operation for the N members of a calibration sample and array this data as in (15).

Performance Variables of the Calibration Sample

n	X(n1)	X(n2)	X(n3)
1			
2			
•			
•			
•			
N			

(15)

From the distributions in the columns of (15), the P(ni) of (3) can be found and arranged as (16).

Percentile of the Performance Variables

n	P(n1)	P(n2)	P(n3)
1			
2			
•			
•			
•			
N			

(16)

For each member of the calibration sample, the L(nij) term of (10) can be found and T(ij), corresponding to (18), can be found as (17).

$$T(ij) = \sum_{n}^{N} L(nij) \qquad (17)$$

$$T(ij) = \begin{array}{c|ccc} i \backslash j & 1 & 2 & I=3 \\ \hline 1 & -- & \sum_n L(n12) & \sum_n L(n13) \\ 2 & \sum_n L(n21) & -- & \sum_n L(n23) \\ I=3 & \sum_n L(n31) & \sum_n L(n32) & -- \end{array} \qquad (18)$$

The T(ij) column sums are estimates of $NI\delta(i)$ if we assume (19)

$$\sum_{i}^{I} \hat{\delta}(i) = 0. \qquad (19)$$

This allows us to estimate $\delta(i)$ by the sum of the T(ij) columns as

$$\hat{\delta}(i) = \sum_{i \neq j} \sum_{n} L(nij)/I \cdot N, \text{ for each } j \text{ and } i \neq j. \qquad (20)$$

Therefore, by (12) we can estimate the comparison (11) and by (14) we can estimate capability, given the difficulties that can be estimated from (20). Note that performers need not perform each of the I operations for their capabilities to be evaluated.

The above estimates are predicated on our ability to estimate the P(ni) and Q(ni) of (3) and (4). This may be accomplished by using (15) and standard statistical methods of fitting data to particular statistical distributions, remembering to include a test for goodness of fit of the observed data to the theoretical distribution.

Example of Application of PM by Method II

As an example of Model II analysis, consider PM for the "job" called Morse Code operation. To make the analysis simple, we assume that the calibration and evaluation samples consist of the same N = 5 persons. Each of these persons has provided I = 3 observations of their performance. These are:

X(ni) = 300 minus the time in seconds required by operator (n) to send a given message.

X(n2) = the number of words sent/20 by operator (n) in 1 minute. (21)

X(n3) = 8 minus the number of errors made by operator (n) during his reception of a given message.

The evaluator has determined from analysis of a large set of records that the X(ni) observations have the following distributions:

X(n1) are normally distributed with $\mu = 101.84$ and $\sigma = 19.71$.
X(n2) have a beta distribution with parameters p=5 and q=3. (22)
X(n3) have a Poisson distribution with parameter m=2.0.

The observed X(ni) data are given below for N = 5, together with their percentiles P(ni). For this analysis, it is important to scale the X(ni) so that increasing (decreasing) magnitude on each variable signifies increasing (decreasing) capability on the job.

n	X(n1)	P(n1)	X(n2)	P(n2)	X(n3)	P(n3)	
1	76	.050	.31	.033	2	.677	
2	152	.999	.70	.647	8	.9999	
3	103	.529	.64	.509	5	.983	(23)
4	147	.998	.83	.899	7	.999	
5	66	.011	.22	.007	3	.857	

Having determined the P(ni) values from knowledge of the X(ni) values and their distributions, we use (9) to compute (10) and array the L(nij) as T(ij) as follows:

$$T(ij) = \begin{array}{c|ccc} i \backslash j & 1 & 2 & 3 \\ \hline 1 & -- & \sum_{n} L(n12) & \sum_{n} L(n13) \\ 2 & \sum_{n} L(n21) & -- & \sum_{n} L(n23) \\ 3 & \sum_{n} L(31) & \sum_{n} l(n32) & -- \end{array} \qquad (24)$$

For the preceeding data:

$$\sum_n L(n12) = -\sum_n L(n21) = 11.297$$

$$\sum_n L(n13) = -\sum_n L(n31) = -16.918 \quad (25)$$

$$\sum_n L(n23) = -\sum_n L(n32) = -28.210$$

These results give T(ij) as:

$$T(ij) = \begin{bmatrix} -- & 11.297 & -16.918 \\ -11.297 & -- & -28.210 \\ 16.918 & 28.210 & -- \end{bmatrix} \quad (26)$$

Adding the columns of T(ij), we get the vector T(i) as:

$$T(i) = 5.621,\ 39{,}507,\ -45.128\ , \quad (27)$$

and, dividing T(i) by NI = 15, we get the estimated operation difficulties:

$$d(i) = \hat{\delta}(i) = 0.374,\ 2.634,\ -3.008\ . \quad (28)$$

Thus, relatively, operation (3) is easy, operation (2) is hard, and operation (1) is of intermediate difficulty. Note also that

$$\Sigma d(i) = 0, \quad (29)$$

satisfying condition (19) for this job.

Our final arithmetical task is to estimate the α(n) for these operators. This is simply done, for, by (6) and (11):

$$L(ni) = \ln\left[\frac{P(ni)}{Q(ni)}\right] = t(ni) = \alpha(n) - \delta(i), \quad (30)$$

$$\sum_i L(ni) = I\alpha(n) - \sum_n \delta(i), \quad (31)$$

but by (19) we estimate α(n) as a(n) with

$$a(n) = \left[\sum_i L(ni)\right]/I. \quad (32)$$

The L(ni) and a(n) for this example are:

n	L(n1)	L(n2)	L(n3)	ΣL(ni)	a(n)	
1	-2.944	-3.378	0.740	-5.582	-1.861	
2	6.907	0.606	9.210	16.613	5.574	
3	0.116	0.036	4.057	4.209	1.403	
4	6.213	2.186	6.907	15.306	5.102	
5	-4.499	-4.995	1.791	-7.663	-2.554	(33)

We note that operators (2) and (4) have the most capability for the Morse Code operator job, whereas operators (1) and (5) have the least. Operator (3) is intermediate to these two sets as measured by our PM Model II.

Examples of Application of PM by Method I

The following illustration is an application of PM Model I to intelligence testing. Although this application relates to a classical psychometric situation and problem, the model applies equally well to situations involving the measurement of performance operations. The common concept of these two situations is that operation performance quality is the observational variable. Quality is classified in one of K categories for each operation of a "job." This may be accomplished either subjectively by the evaluator or his agent, by objective evidence, or both. The quality categories are the same for each operation. The analytic details of this PM model are given in Moonan (1979). For simplicity, a sample case where K=3 (i.e., we have K=3 performance quality categories that we label "outstanding," "adequate," and "poor") was selected. The standards for responses of these qualities are given objective definitions. The I=3 items for the miniature intelligence test used in this research have been selected from various technical sources. The item responses are abbreviated briefly as (+ 0 -). The items were selected so that the "test" that they comprise measures the latent trait called "reasoning ability." The first item is a puzzle presented by Fisher (1937, p. 86). The puzzle's solution requires the construction of 4x4 Latin Square experimental designs. The second item is a chess problem from Sweeney and Barclay (1972, p. 1). The chess problem requires the player (White) to mate the Black King in two moves. The third item is a problem in elementary probability theory given by Rubenstein (1975, p. 185). The solution requires, in effect, the application of Bayes' theorem. The solutions to these items are shown below:

1. The Welsh lawyer is a 45-year-old married socialist.
2. Q-h3, K-e4 forced; R-c4, MATE. (34)
3. Pr [Dem | Vote Rep] = .25.

Suppose that these three items have each been attempted by N = 300 members of a calibration sample and the frequencies F [i,j,ℓ,m], by which the sample members gave responses of quality ℓ to item i <u>and</u> responses of quality m to item j, are portrayed as shown below:

		j=1			j=2			j=3		
i	ℓ \ m	-	0	+	-	0	+	-	0	+
	-	136	-	-	22	40	74	2	16	118
1	0	-	96	-	16	28	52	2	11	83
	+	-	-	68	11	20	37	1	8	59
	-	22	16	11	49	-	-	1	6	42
2	0	40	28	20	-	88	-	1	10	77
	+	74	52	37	-	-	163	3	19	141
	-	2	2	1	1	1	3	5	-	-
3	0	16	11	8	6	10	19	-	35	-
	+	118	83	59	42	77	141	-	-	260

(35)

We perform our analysis for $C_2^K = K(K-1)/2 = C_2^3 = 3$ pairs of response categories (ℓ=1, m=2), (ℓ=1, m=3), and (ℓ=2, m=3). For the first pair, we extract (36) from (35), giving:

F [i,j,1,2] =

i \ j	1	2	3
1	-	40	16
2	16	-	6
3	2	1	-

(36)

From (35), we construct the ratios:

R [i,j,1,2] = F [i,j,1,2]/F [i,j,1,2] as (38). (37)

R [i,j,1,2] =

i \ j	1	2	3
1	-	2.500	8.000
2	.400	-	6.000
3	.125	.167	-

(38)

We also obtain, in (40), the natural logarithms of the R(i,j,1,2); thus,

lnR [i,j,1,2] = L [i,j,1,2] (39)

L [i,j,1,2] = i \ j	1	2	3	
1	-	0.916	2.079	
2	-0.916	-	1.792	(40)
3	-2.079	-1.792	-	

Adding each column of L [i,j,1,2], we get T(i) as:

$$T(i) = [-2.995 \quad -.876 \quad 3.871]. \tag{41}$$

Finally, dividing T(i) by -3/2 of (42), we get estimates d(i) of the operation difficulties, $\delta(i)$ of (43),

$$\Delta(\ell,n) = -I \;\; \Theta(3) - \Theta(2) \;\; , = -3(1 - 1/2) = -3/2, \tag{42}$$

where, for this example,

$$\begin{aligned} I &= 3 \\ \Theta(3) &= 1 \\ \Theta(2) &= 1/2 \\ \Theta(1) &= 0 \\ d(i) &= [+2.000 \quad +.584 \quad -2.581]. \end{aligned} \tag{43}$$

The Θ values are scale points on the response distribution. This process is repeated for response pairs (ℓ=1, m=3) and (ℓ=2, m=3). This effort yields the three sets of $\delta(i)$ estimates that we average to get the final d(i)s.

(ℓ,m) \ i	1	2	3	
(1,2)	2.000	.584	-2.581	
(1,3)	2.226	.244	-2.470	(44)
(2,3)	2.197	.295	-2.492	
AVG	2.141	.374	-2.514	

The capability estimates for members of an evaluation sample are found by a process known as expected value estimation (EVE), as shown in Moonan (1979). Moonan also found that Model I has the important property of "specific objectivity."

CONCLUSIONS

1. Model I is a performance model that can be efficiently used to measure performance by means of test items. It is simple to understand and to use. Also, the model has the property of specific objectivity.

2. Model II is a performance model that can be effectively used to measure performance of persons in jobs requiring a variety of operations, each possessing its own form of observational variable distribution.

3. Job performance measurement models are conveniently developed as functions of the comparison between the performer's capability and the difficulties of the operations to be performed on the job.

RECOMMENDATIONS

1. Model I should be used to assess abilities of persons who are to be evaluated by tests in which items are scored with dichotomous responses.

2. Model II should be used to assess capabilities of persons to perform jobs whose operations are evaluated by means of variables with different distributional forms.

REFERENCES

Choppin, B. Item bank using sample-free calibration. Nature, 1968, 219, No. 5156.

Fisher, R. A. Design of experiments. London: Oliver and Boyd, 1937.

Moonan, W. J. Implementing the polychotomous response measurement model (FACT Document No. 64). San Diego: Navy Personnel Research and Devlopment Center, 1979.

Panos, R. Measurement--its sense and nonsense. Proceedings of 1969 Military Testing Association Conference, New York, 1969.

Rasch, G. Probability models for some intelligence and attainment tests. Copenhagen, Denmark: Danish Institute for Educational Research, 1960.

Rubenstein, M. Patterns of problem solving. New York: Prentice-Hall, 1975.

Sweeney, T., & Barclay, W. Beauty is where you find it. Brooklyn, NY: Pageant-Poseidon Ltd., 1972.

Wright, B. D. Sample-free test calibration and person measurement. Proceedings of 1967 Invitational Conference on Testing Problems, ETS, Princeton, NJ, 1967.

Wright, B. D., & Mead, R. J. BISCAL: Calibrating items and scales with the Rasch model (Res. Memo. No. 23A). Chicago: Statistical Laboratory, Department of Education, University of Chicago, 1978.

DISTRIBUTION LIST

Chief of Naval Operations (OP-964D), (OP-987H)
Chief of Naval Material (NMAT 0722), (NMAT 08L)
Chief of Naval Research (Code 450) (2), (Code 458) (2)
Chief of Naval Education and Training (02), (015), (N-2), (N-5), (N-9)
Chief of Naval Technical Training (016)
Commandant of the Marine Corps (Manpower Utilization)
Commanding Officer, Naval Health Research Center
Commanding Officer, Naval Training Equipment Center (Technical Library)
Commanding Officer, U.S. Army Research Institute for the Behavioral and Social Sciences
Commander, U.S. Army Soldier Support Center, Fort Benjamin Harrison (Human Dimensions Division)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base
Commander, Air Force Manpower and Personnel Center, Randolph Air Force Base
Defense Technical Information Center (DDA) (12)

ILMED

8