REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) TN 88-39	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Kevin Murphy 6b OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Navy Personnel Research and Development Center
6c. ADDRESS (City, State, and ZIP Code) Colorado State University Fort Collins, CO 80523	7b. ADDRESS (City, State, and ZIP Code) San Diego, California 92152-6800
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAAG 29-81-D-0100, Deliver Order 2065
8c. ADDRESS (City, State, and ZIP Code) Washington, D.C. 22217-5000	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62763N; PROJECT NO. R63521; TASK NO. 804; WORK UNIT ACCESSION NO 040

11 TITLE (Include Security Classification)
Dimension of Job Performance

12. PERSONAL AUTHOR(S)
Kevin Murphy and Leonard Kroeker

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Feb 86 TO May 86	1988 April	37

16 SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Job Performance, Performance Measures, Performance Dimensions
05	09		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
This report describes and applies a method of determining the dimensions that make up the the performance of work throughout the Navy's enlisted ratings and grades. The dimensions of work performance are defined within a hierarchy from micro to macro and the relationships between such dimensions and input variables such as ability, experience, and motivation are described. Key inclusions and omissions from current research within DoD on the prediction of job performance are identified.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL ROBERT F. MORRISON	22b. TELEPHONE (Include Area Code) (619) 553-9256 / 22c. OFFICE SYMBOL 62

Navy Personnel Research and Development Center
San Diego, CA 92152-6800 | TN 88-39 | April 1988

Dimensions of Job Performance

DEPARTMENT OF THE NAVY
NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA 92152-6800

3900
Ser 62/363
26 APR 1988

From: Commanding Officer, Navy Personnel Research and Development Center

Subj: **DIMENSIONS OF JOB PERFORMANCE**

Encl: (1) NPRDC TN 88-39

1. This research and development was conducted within the exploratory development project RF63-521-804 (Manpower and Personnel Technology), work unit 040-03.01 (Dimensions of Job Performance). The purpose of the work unit is to define general dimensions that describe the global construct of human performance at work and to identify measures of such dimensions. Such dimensions will provide a framework for estimating how effective a single measure may be in predicting job performance. This exploratory development was conducted under contract DAAG 29-81-D-0100, delivery order 1907.

2. Enclosure (1) is the second in a series produced under this work unit. The previous report described factors that made it inappropriate to try to use personnel record data to develop a surrogate measure of job performance that would generalize across ratings and grades.

3. Point of contact at NAVPERSRANDCEN is Dr. Robert F. Morrison, AUTOVON 553-9256 or Commercial 553-9256. Comments are welcome.

JOHN J. PASS
By direction

Distribution:
Deputy Assistant Secretary of the Navy (Manpower)
Chief of Naval Operations, Assistant for Planning and Technology Development (OP-01B2); Director, Military Personnel Policy Division (OP-135L)
Technology Area Manager, Office of Naval Technology (Code 222)
Chief, U.S. ARI-USAREUR
Commander, Air Force Human Resources Laboratory, Manpower and Personnel Division (AFHRL/MO), Brooks Air Force Base, TX
Center for Naval Analysis
Superintendent, Naval Postgraduate School
Defense Technical Information Center (2)

NPRDC TN 88-39

April 1988

Dimensions of Job Performance

Kevin R. Murphy, Ph.D.
Colorado State University
and
Leonard P. Kroeker, Ph.D.
Navy Personnel Research and Development Center

Reviewed by
Robert F. Morrison, Ph.D.
Gerald J. Laabs, Ph.D.

Approved and released by
John J. Pass, Ph.D.

Navy Personnel Research and Development Center
San Diego, California 92152-6800

SUMMARY

Problem

Each of the services is currently involved in developing operational measures of job performance in the enlisted ranks. The validity of these measures depends in part on the extent to which they cover the major dimensions of the job performance domain. However, neither the dimensions of job performance nor a practical method of determining those dimensions have been identified to date. The development and application of a method for identifying the dimensions of job performance is there fore a prerequisite for determining the validity of operational performance measures. Identification of the dimensions of job performance will provide the first step in assessing the utility of several performance measurement technologies, as well as assessing the extent to which specific measures cover the domain of job performance.

Purpose

The purpose of this report is to describe and apply a method of determining the dimensions of job performance in the enlisted ranks of the navy. Job performance dimensions are defined and are placed in a framework that (a) describes the hierarchical relationships among performance dimensions, and (b) describes the relationships between input variables such as ability, experience, and motivation and these performance dimensions. Initial estimates of the validity of several classes of performance measure are obtained by relating those measures to each of the dimensions of the performance domain.

Approach

Several potential methods of identifying performance dimensions are reviewed, and a method that involves specifying relevant goals and behaviors that relate to the attainment of those goals is suggested. A set of goals that includes readiness, technical proficiency, teamwork, and task accomplishment is described, and nine behavioral dimensions are derived that relate to those goals. On the basis of related research literature, these dimensions are arranged in a hierarchical framework that relates each performance dimension to several input variables. On the basis of definitions of each performance dimension and of eight different performance measures, operational measures are related to both individual performance dimensions and to groups of dimensions.

The definitions of each performance dimension are used to relate individual dimensions and groups of dimensions to input variables. On the basis of these relationships, predictions are made regarding the probable outcomes of criteria-related validity studies. The effects of contextual moderators on the relationships described above are also specified.

Results

A framework describing each of the major dimensions of job performance, the relations among these dimensions, and the relationships between input variables and performance dimension is developed and presented in graphic form. The rationale for the proposed relationships, along with the set of variables that affect these relationships, is described in detail.

Conclusions

The method described here provides a practical basis for specifying the dimensions of job performance. These dimensions include both job proficiency and task accomplishment, both of which may be adequately measured by methods currently being developed by several of the services. However, other important dimensions, particularly those interpersonal and work-avoidance behaviors, are not adequately covered by current measurement technologies.

Input variables such as ability and motivation are clearly related to several of the performance dimensions. However, no single input variable has the same relationship to all performance dimensions. One consequence is that the true relationship between input variables and performance measures may vary considerably from measure to measure. The performance framework described in this report provides a useful basis for assessing both the validity of selection and placement tests for predicting job performance, and the validity of specific operational indices in measuring job performance.

Recommendations

Empirical research is needed to test the conceptual models and theoretical predictions described in this report. The model employed here suggests that no single measure will provide adequate coverage of the job performance domain; multiple performance measures employing distinct measurement technologies are recommended. Taxonomic research on the dimensions of job performance is also recommended. Additional dimensions or alternative methods of classifying performance dimensions may emerge from a more rigorous research effort. Finally, research on dynamic relations among performance dimensions and between inputs and performance is recommended. It is likely that the network of relationships among inputs and performance changes as new recruits gain knowledge, skills, and experience; at present, little is known about the nature of these changes.

CONTENTS

CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Problem

Following a Congressional mandate, each of the Services is currently involved in developing measures of job performance for enlisted personnel.[1] These measures range from behaviorally anchored ratings to job simulations; each of these measures is being evaluated as a practical and economical substitute for the mandated bench mark of hands-on performance. It is anticipated that these new measures will serve several functions in the military personnel systems, particularly in the validation of selection and placement tests. The validity of these performance measures is there fore a critical concern.

The validity of performance measures can be defined in several ways. One way is to determine whether these performance measures adequately cover the domain of job performance, and if not, what sorts of additional measures would be needed to cover the domain. To make this determination, the major dimensions of job performance in the enlisted ranks must be identified and defined in sufficient detail to allow the matching of performance measures to specific dimensions of the job performance domain.

The dimensions of job performance in the enlisted ranks of the Navy (or of the other Services) have not been identified to date. At present, no standard methods exist that can be used to identify the dimensions of job performance. The development and application of a method for identifying job performance dimensions is therefore a prerequisite for determining the extent to which operational performance measures cover the domain of job performance.

Purpose

The purpose of this report is to describe a method for determining the dimensions of job performance in the enlisted ranks of the Navy. This method is applied, and performance dimensions are derived. A framework describing the performance domain is then constructed that (a) describes the hierarchical relationship among performance dimensions, (b) describes the relationships between input variables such as ability, experience, and motivation, and these performance dimensions, and (c) accounts for the effects of situational constraints on the relationships described in (b). This framework can then be used to relate operational performance measures to specific performance dimensions or to sets of performance dimensions. It is also possible to derive predictions regarding the criterion-related validity of ability, experience, and personality measures in predicting both general performance dimensions and specific performance measures.

By describing in detail the dimensions that underlie the domain of job performance, it becomes possible to describe the relationships among diverse research efforts currently underway in the Navy and in other Services. Thus, the framework developed here not only serves the immediate need of assessing the validity of operational performance measures, but also provides a standard frame of reference for integrating a broad range of research efforts aimed at predicting, improving, and measuring job performance in the enlisted ranks.

---

[1]Several of these efforts are related to Project A, which seeks to develop an integrated, state-of-the-art system of personnel assessment and classification. Feedback from several Project A researchers is gratefully acknowledged.

DEFINING THE PERFORMANCE DOMAIN

Although job performance is one of the central concepts in personnel administration, the domain of job performance is poorly defined and poorly understood. There have been few attempts to define in any detail the boundaries and structure of that domain, or to identify the major elements of that domain. Thus, there is typically no clear definition of what is meant by "job performance." This problem is clearly manifest in research on task performance, job performance, and productivity; in all three areas, there is little agreement concerning the variables that underlie and define the meaning of "performance." This is most obvious in research on productivity. Three strategies seem to dominate in defining productivity: (1) to leave productivity undefined (cf., Sutermeister, 1976), (2) to admit that the concept is ill-defined and propose a vague definition (cf., Muckler, 1982), or (3) to propose several definitions, each of which applies in different circumstances (cf., Guzzo, Jette, & Katzell, 1985). Thus, it is difficult to agree whether productivity is going up or down (and if so, by how much), whether productivity enhancement programs really work, and most important, what the consequences would be if productivity levels were to change substantially.

In research on task performance, definitional problems are not so readily apparent. Yet, as the work of Fleishman and his associates has illustrated, the classification and definition of task performance is exceedingly complex (Fleishman & Quaintance, 1984). Part of the problem is that it is difficult to define the boundaries of complex tasks. That is, it is often difficult to delimit task behavior from non-task behavior; this problem becomes especially acute when task behavior includes mental as well as physical labor. In addition, performance on several common tasks is both complex and multidimensional, especially when considered in a real-time context in which several tasks must be attended to. Thus, it can be difficult to determine precise links between an individual's behavior and his or her progress in completing a specific task.

There is an extensive literature dealing with the prediction and measurement of job performance. For example, Landy and his colleagues have reviewed research on performance rating and have suggested several innovative directions for further research and application (Landy & Farr, 1980, 1983; Landy, Zedeck, & Cleveland, 1983). In the area of prediction, Schmidt, Hunter, and their colleagues have reviewed hundreds of studies that suggest that ability tests provide valid predictions of job performance (Hunter & Hunter, 1984; Pearlman, Schmidt & Hunter, 1980; Schmidt & Hunter, 1977, 1981). Both of these literatures are notable for the fact that job performance, which is the central concern of the research reviewed, is rarely if ever defined. As a result, we know a great deal about predicting future performance, and also know a great deal about measuring past performance, but know little about what we have predicted or measured. Campbell (1983) notes that we have been quite successful in modeling and defining different parts or aspects of performance, but that little effort has gone into defining the overall performance domain.

Job Performance vs. Task Performance

One possible explanation for this lack of attention directed to defining job performance is that job performance is typically equated with task performance. That is, it is often assumed that job performance can be adequately defined in terms of the incumbent's success in carrying out the tasks that are included in a set of occupational standards. From this perspective, the growing body of research dealing with the determinants and the nature of task performance (cf., Fleishman & Quaintance, 1984; Salvendy & Seymour, 1973) provides the key to understanding job performance. If job

performance is equated with task performance, research priorities in this field should be directed at developing better job descriptions and at aggregating measures of performance on several separate tasks into an overall measure of job performance.

There are several reasons to believe that job performance cannot be equated with task performance. First, most observations of work behavior confirm the common perception that workers spend relatively little time performing what would be regarded as tasks. For example, Bialek, Zapf, & McGuire (1977) reported that enlisted infantrymen spent less than half of their work time performing the technical tasks for which they had been trained; in many cases, only a small proportion of an enlisted man's time was in any way devoted to accomplishing the tasks specified in his or her job description. Campbell, Dunnette, Lawler, & Weick (1970) noted similar patterns for managers. A substantial part of the manager's day is spent doing things which cannot be unambiguously linked to the accomplishment of specific tasks. The fact that most people's work time is not devoted solely to tasks has serious implications for several aspects of criterion development. For example, many methods of job analysis are based explicitly on the assumption that most, if not all, of a worker's time is spent working on identifiable tasks (e.g., the Air Force's task analysis system. See Christal, 1974). Since much of the work day is spent doing something outside of the typical domain of tasks, indices of the percentage of time spent on different tasks present a warped view of the activities actually carried out by workers. Unless you are willing to ignore much of what a person does at work, it is difficult to equate job performance with task performance.

Second, many of the performance evaluation systems currently in use in the civilian and military sectors include specific measures or indices that are only tangentially related to task performance. Examples include measures of absenteeism and turnover, as well as supervisory ratings of broad traits such as dependability or motivation. Admittedly, the use of such measures does not prove that the job performance domain is broader than the task performance domain; it is possible that these measures, although widely used, are invalid. Nevertheless, the widespread use of measures that do not relate directly or solely to the accomplishment of tasks does suggest that the job performance domain is perceived to be broader than the domain of task performance.

A third argument for assuming that job performance cannot be defined solely in terms of task performance is that job performance must be defined over longer time periods and in relation to more organizational units than is true for task performance. This can be seen most clearly by considering the case of a machinist's mate who successfully accomplishes his tasks by depleting all available reserves of material and by diverting resources from other work units. Although the individual successfully achieved his main task, the long-term implications of such task performance are clearly not favorable. In many jobs, it would not be difficult to provide examples of individuals who successfully completed most tasks, but whose performance was judged to be low.

Elements of the Performance Domain

In defining the domain of job performance, two questions must be answered. First, what are the boundaries of this domain? In other words, what is the "stuff" of job performance. Second, how is this domain organized? Both of these questions have been addressed in part by several researchers.

The fundamental question in determining the content of the job performance domain is whether performance should be defined in terms of behavior or in terms of the results of behavior (James, 1973; Smith, 1976). From the organization's point of view, there is a

strong temptation to define performance in terms of results. There are, however, several reasons for defining performance in terms of behavior rather than in terms of results. First, an exclusive emphasis on results is likely to lead to behaviors that are dysfunctional for the organization. Landy and Farr (1983) note that if performance is defined exclusively in terms of countable outcomes *e.e., results), incumbents will be strongly motivated to maximize those outcomes at the expense of other activities (e.g., maintenance, planning, conservation) that are vital to the organization. Second, results are more complexly determined than behaviors, in that results are a joint function of what the person does and the situation in which he or she does it. Thus, until we more fully understand the domain of job behaviors, it may not be possible to fully understand the joint effects of behaviors and situations. The consideration outlined above suggests that job performance should be defined as a domain of behaviors which occur on the job, or in conjunction with the job. As will be noted later, "job behavior" can be broadly defined to include highly specific acts as well as general patterns of behavior (i.e., traits). Several types of behaviors might be included in this domain; research aimed at determining the dimensions of criteria is relevant in defining the boundaries of that behavioral domain.

Research on the dimensions that underlie various measures of job performance has led to two general conclusions. First, it is widely agreed that performance is multidimensional (Bass, 1982; Dunnette, 1963; Pickle & Friedlander, 1967; Ronan & Prien, 1966; Seashore, 1975). Second, there appears to be no strong general factor underlying analyses of the most common performance measures (James, 1973; Smith, 1976). Taken together, these two conclusions might lead to the further conclusion that concepts such as "overall performance" or "general level of performance" are potentially meaningless. However, these concepts are widely used in practice (Landy & Farr, 1983), are thought to be necessary for administrative purposes (Schmidt & Kaplan, 1971), and underlie the vast body of literature tying personnel and selection to outcomes such as productivity (Hunter & Hunter, 1984; Hunter & Schmidt, 1982; Murphy, 1986; Schmidt & Hunter, 1977, 1981; Schmidt, Hunter, McKenzie, & Muldrow, 1979). In part, the gap between the body of research which suggests that there may be no such thing as "overall performance" and the body of research which treats "overall performance" as its central concern can be traced to differences in method and focus. The failure to find a strong general factor may say more about the shortcomings of the methods used to define the performance domain than about the domain itself.

Methods of Determining Performance Dimensions

Four methods could be applied in determining the behavioral dimensions that define the domain of job performance. First, you could obtain a representative sample of the different measures that are commonly used in measuring job performance, and apply factor analysis or some related techniques to the intercorrelations among these measures. Rush (1953) applied this strategy in studying sales criteria, and suggested that the underlying dimensions of sales success could be defined as (1) Objective Achievement, (2) Learning Aptitude, (3) General Reputation, and (4) Sales Technique. Ronan (1963); Richards, Taylor, Price, and Jacobsen (1965); and Turner (1960) reported similar analyses, yielding from 4 to 29 factors, depending on the job category and the criterion measures contained in each study.

The factor-analytic approach is basically inductive, in that it attempts to discover the nature of performance by analyzing various measures of performance. The difficulty with this approach is that it involves two untestable, and probably untenable, assumptions. First, one must assume some level of content or construct validity for each of the measures included in the analysis. If these measures will be used to define the dimensions

of performance, it must be assumed that they all have something to do with performance. More important, this method assumes that the set of performance dimensions studied spans the performance domain; dimensions such as work avoidance will emerge if absenteeism measures are included, but will not emerge if data from personnel files are not included in the analysis. Thus, the decision to either include or exclude a specific measure from your study (e.g., a particular measure of absenteeism) represents an a priori decision about the boundaries and structure of the performance domain. If you knew which measures were and were not included in the content domain, and also knew which measures must be included to obtain a representative sample of that domain, it is unlikely that you would need to do a factor analysis in order to determine the nature and the structure of the domain.

A very different method of determining the dimensions of job performance is to observe and analyze incumbents' behavior—in other words, to conduct some sort of job analysis. For example, Hemphill (1959, 1960) suggested that performance as an executive involved dimensions such as long-range planning, exercising power, and community involvement. McCormick and his colleagues have developed structured job analysis questionnaires that assess the behavioral elements common to performance in a wide variety of jobs (McCormick, 1979; McCormick, Jeanneret, & Mecham, 1972). Factor analyses of these questionnaires would certainly provide useful information regarding the dimensions of job performance. The problem with the job analytic approach is that it is difficult to obtain meaningful observations without providing some structure, and that the structure of the measurement instrument itself will have a strong impact on your eventual conclusions regarding the dimensions of job performance. For example, questionnaires used to analyze Air Force jobs concentrate exclusively on discrete tasks (Christal, 1974). The dimensions which emerge from analyses of these questionnaires will therefore be task dimensions. The PAQ samples a broader array of behaviors, but concentrates heavily on issues such as the use of tools and machines. The content of the PAQ questions will affect the pattern of results coming from an analysis of PAQ responses. In a general sense, the job analysts' decision of what to observe and how to quantify those observations will have some impact, and the dimensions uncovered through job analytic methods cannot be regarded as pure dimensions of job performance. Rather, they represent an interaction between the structure of the domain and the structure of the instrument chosen to analyze the domain. As was true with factor analyses of performance measures, dimensions based on job analysis data may say more about the measures analyzed (and not analyzed) than about the domain. Unless an adequate sampling of the domain can be guaranteed prior to the analysis, these methods may be insufficient for revealing the dimensions of job performance.

The methods of human factors engineering present another approach to defining the dimensions of job performance (Chapanis, 1976; Campbell, 1983). Rather than using statistical analyses of performance measures or observations to uncover these dimensions, it might be possible to analyze the physical and cognitive and psychomotor demands of tasks, and to derive dimensions on the basis of task demands (Fleishman & Quaintance, 1984). There are, however, two limitations to this approach. First, this approach works well with simple tasks, but is difficult to apply to complex, poorly defined tasks (Campbell, 1983). Second, this approach limits its focus to the domain of task performance. It would be difficult to apply this approach to work behaviors, such as absenteeism, that are believed to be important aspects of performance.

A fourth approach to defining performance dimensions is the construct-oriented strategy. Although this approach has been widely advocated (James, 1973); Smith, 1976), it has rarely been applied. Here the task of defining the dimensions of job performance

could be regarded as an exercise in construct explication (Nunnally, 1978). Using this approach, performance dimensions are defined rather than discovered; the set of dimensions which are decided upon depends on the definition employed for the construct "performance."

The construct-oriented model is appealing for several reasons. First, a construct-oriented strategy has long been recognized as the optimum method for assessing the validity of performance measures (Guion, 1980; James, 1973). The same model that is so highly recommended for the specific problem of assessing individual measures is obviously relevant to the more general problem of defining performance dimensions. Second, as noted above, the construct-oriented approach provides a way of avoiding the trap implicit in the inductive approach--that performance dimensions can be discovered, if only the right methods are applied. The construct-oriented approach demands that these dimensions be defined as part of the process of construct explication. Furthermore, the construct-oriented approach does not place arbitrary statistical limitations on the nature of the performance dimensions. Factor analysis produces dimensions that are orthogonal, and that account, in descending order, for the greatest proportion of the common variance, etc. While these statistical features are useful, there is rarely any reason to suspect that the underlying dimensions of performance, or of other behavioral constructs, do show any of these characteristics.[2] Dimensions that are defined rather than derived need not reflect these potentially arbitrary statistical limitations. Finally, the construct-oriented strategy makes explicit the assumption that performance is a construct, and that the ultimate definition of performance dimensions depends entirely on our conceptual definition of performance. In the section that follows, a construct-oriented strategy is applied to define performance, to define the dimensions of performance, and to specify the interrelations among the dimensions of job performance.

DEFINING PERFORMANCE DIMENSIONS

Rather than using an inductive approach to discover the dimensions of job performance, one might define those dimensions in terms of their logical relations to the construct of job performance. The preceding section suggests that these dimensions should be behavioral dimensions, but says little about what behaviors should or should not be included in a definition of performance. Astin's (1964) definition of conceptual criteria provide a useful tool for defining the behavioral elements of job performance.

Astin (1964) noted that in developing criteria, we must identify the relevant goals of the sponsor or the measurer. In the context of work, the relevant goals of the organization would include both short-term goals, such as the successful completion of specific tasks, as well as long-term goals, such as the maintenance of effective relations between work group, departments, etc. According to Astin (1964), the conceptual criterion is nothing more than a verbal abstraction of the relevant goals or the outcomes desired; the set of possible criterion measures would include any observable index or state that is judged relevant to the conceptual criterion. If we substitute "performance construct" for "conceptual criterion," and "performance dimensions" for "criterion measures," it becomes possible to define the domain of performance and to indicate the set of behaviors that are included in that domain. The performance domain is defined here as the set of behaviors that are relevant to the goals of the organization or the organizational unit in which a person works.

---

[2]Not all methods of confirmatory factor analysis involve these assumptions (Bentler & Weeks, 1980; Joreskog, 1969; Long, 1983). However, any method of factor analysis imposes some statistical structure on the dimensions obtained.

In order to specify the range of behaviors that define performance, one would have to know the relevant goals of the organization. Note here that the global set of goals that define the overall effectiveness of the organization (e.g., maximize after-tax profits) are not as relevant as the set of goals which are defined for an incumbent in a specific position within the organization. That is, the organization defines a set of goals to be met by incumbent in each job, and the relevant goals may vary considerably from job to job, or across different levels in the organization. This is particularly true when specific task goals are considered, but will also be true for several non-task goals. For example, one goal which is likely to be broadly relevant, but which is not tied to any specific task, is that incumbents must maintain effective interpersonal relations with their co-workers and with other organizational members with whom they interact. The type and extent of these interpersonal contacts will vary across departments and across levels of management. One might infer that the skills and behaviors which contribute to successful maintenance of interpersonal relations will vary from job to job. Nevertheless, within the great majority of jobs, this general class of behaviors is likely to represent one aspect of effective performance.

Since the relevant goals of organizations and of organizational units differ, it may not be possible to draw up a completely general definition of the dimensions of job performance. Nevertheless, there are enough broad similarities in organizational goals and in job demands to justify a framework that is relevant for defining performance in a large class of jobs. For example, the performance dimensions that are relevant for describing skilled craft jobs (e.g., electrician, plumber) would vary in some of their specifics, but would probably show considerable communality.

Level of Analysis

The goals and the performance dimensions identified for large, diverse organizations are likely to be for the broad and general; an analysis of a single job, or a homogeneous set of jobs, might yield performance dimensions that were considerably more specific and refined. It is not clear that there is any advantage to deciding what level of analysis is the "correct" level for defining the performance domain. Different levels of analysis may be useful for different purposes. Thus, it may be inappropriate to ask what are the dimensions of performance. It is clear, however, that the level of analysis will affect the sorts of goals that are defined as relevant, and there fore will affect the specification of the behaviors that are most relevant to those goals (i.e., performance dimensions). First, there is evidence that the goals of an organization are generally different from (and more specific than) those that the organization's official or implicit mission statement (Rice, 1963). That is, organizational goals cannot be specified in any detail simply by knowing the general mission or the organization (e.g., profit, service, regulation, protection). Second, the goals of units within organizations are not necessarily related to organizational goals (Hunt, 1976; Trist, Higgins, Murray, & Pollack, 1963). For example, the owners of a mine might set goals that maximize production, given a certain margin for safety, but work group goals might maximize safety, given a certain margin for performance. These two sets of goals will yield different ideas of what performance is, of whether workers are performing well or not, and of the extent, the nature, and the causes of individual differences in job performance.

Deriving Goals

Once the relevant goals of an organization or an organizational unit have been defined, the process of defining performance dimensions is relatively straightforward. The definition of these goals, however, can be a complex matter, especially since the

professed goals of a work unit might be quite different than the goals actually pursued by that unit (i.e., operating goals) (Perrow, 1961).

A sociotechnical systems approach provides one method for deriving performance goals and performance dimensions. This approach assumes that the nature and definition of a job and of the work carried out by an incumbent depends on both the social and the technological organization of work behavior (Emery & Trist, 1960; Herbst, 1974; Rice, 1958; Trist & Bamforth, 1951; Woodward, 1965, 1970). The sociotechnical analysis of work stresses the detailed observation of intact work units, with particular attention given to the interdependence between workers and work groups, the interplay between technology and the methods employed by individuals and work groups to accomplish major tasks, and to the effects of changes in either the social or the technical milieu on the work and on the worker. Other methods that might prove useful in defining goals include the use of surveys, the analysis of the formal structure of the organization, and the analysis of organizational outputs.

The choice of a method for defining relevant goals will depend largely on the level of analysis chosen. There is a positive correlation between the level of generality of the unit of analysis and the level of generality of the goals that are likely to be derived. For example, goals that span an organization as large and diverse as the Navy will necessarily be quite general. Performance goals for a homogeneous subset, such as all machinist mates in the Navy would be more specific. Goals for a particular group, such as the machinist's mates in the Navy would be more specific. Goals for a particular group, such as the machinist's mates assigned to a particular aircraft carrier, would be more specific still.

DERIVING PERFORMANCE DIMENSIONS FOR NAVY ENLISTED OCCUPATIONS

To illustrate the application of a construct-oriented approach to defining the dimensions of performance in a large class of jobs, a framework was developed for describing the domain of performance as an enlisted man or woman in the Navy. This category of jobs includes administrative, technical, and warfare-oriented jobs on both ship and shore. Job titles range from Boatswain's Mate to Data Processing Technician, from Missile Technician to Aviation Storekeeper. Although the specific goals associated with individual jobs vary, there are some general goals that are relevant to almost all enlisted positions. One such goal is readiness. In peacetime, one of the major goals of the Navy is to maintain a high state of readiness of both machines and personnel. This is one of the reasons that the Services place a heavier emphasis on training and practice than is true in the private sector. A second relevant goal is to attain and maintain technical proficiency. Since a large number of enlisted men and women serve for periods of four to six years, it is important that recruits quickly attain the complex technical skills needed to perform many of the jobs on both ship and shore. A third relevant goal is effective teamwork. In highly interdependent systems, such as ships or aviation wings, it is natural to emphasize the performance of units rather than the performance of individuals; an individual who performs his or her task effectively, but who impedes the work of the unit as a whole is likely to be viewed as an ineffective worker. Finally, as is the case in almost any job, successful accomplishment of major job tasks is a relevant goal.

Performance Dimensions

This set of goals can be used to derive global dimensions of work behavior that define performance in the enlisted ranks in the Navy. A set of nine separate performance

dimensions is illustrated in Figure 1.[3] The organization of this domain is somewhat complex; the nine separate elements are divided into six groups, and are described below.

1. Effectiveness in position is affected by four classes of behaviors: (a) Task Performance Behaviors--successful accomplishment of work-related goals and tasks, (b) Destructive/Hazardous Behaviors--that result in or run the clear risk of loss, damage, or productivity setbacks; examples include accident involvement, failure to follow safety procedures, and security violations, (c) Down Time Behaviors--that directly or indirectly lead to significant periods of time in which the enlisted man is not capable of engaging in task performance, or is incapable of performing at his normal level; examples include absenteeism, AWOL, desertion, lateness, work avoidance, alcohol/drug abuse, and court martials, (d) Interpersonal Relations--behaviors that affect the smooth functioning of the work unit and relations between work units; examples include interpersonal conflicts, communication of relevant information, conflict with supervisors, and personal reliability.

Behaviors in categories b, c, and d are likely to have a different functional relationship to job performance than are behaviors in category a. In particular, Destructive/Hazardous and Down Time behaviors serve to limit performance if they are present. The absence of these behaviors, however, does not guarantee good performance. The functional relationship between interpersonal relations and performance will vary according to the job context. In some cases, poor interpersonal relations will hinder job performance, but better-than-average interpersonal relations will not lead to increased performance. Task performance, on the other hand, is likely to have a consistent positive relationship with job performance.

2. Task Performance. The difference between job proficiency, task performance, and job performance is similar to the difference between the tasks one can perform, the tasks one will perform, and the total set of task and non-task behaviors that determine evaluations of effectiveness. Within this framework, task behaviors represent a subset of the relevant job behaviors; task performance criteria are concerned with this subset. Performance as an individual and as a member of a team are considered separately. In particular, interpersonal relations will generally have a greater impact on the former than on the latter.

3. Non-Task Behaviors That Affect Performance. There are several types of behaviors that serve to limit job performance, even when major tasks are carried out effectively. Down Time and Destructive/Hazardous behaviors belong in this group. This class of behaviors is unique in terms of its functional relationship with other types of criteria. These behaviors are also (hopefully) unique in that they have low base rates. One implication is that correlational methods will do a poor job summarizing the effect of these behaviors and the place of this class of behaviors in the total performance domain.

4. Interpersonal Relations. Includes effective communications as well as effectiveness in dealing with others. Interpersonal relations contribute to effective team performance, but may have only a limiting effect on other performance categories.

---

[3] Because of the large number of tables and figures in this report relative to the amount of text, both tables and figures are placed at the end of the report, commencing on page 22.

5. Job Proficiency. Both task performance and job proficiency criteria are concerned with the same subset of behaviors (task behaviors). The difference between the two lies in the conditions under which these tasks are performed. Task performance criteria deal with performance under normal working conditions in which: (a) there are no competing task demands, (b) the subject is motivated to perform at the maximum level and to follow standard procedures, and (c) performance is continually monitored and evaluated in considerable detail.

6. Job-Related Skills and Task Knowledge. In order to attain proficiency, a body of factual and procedural knowledge, together with a set of skills ranging from psychomotor skills to analytic skills, must be mastered. These skills and knowledge bases represent basic components; proficiency is achieved only when these basics are effectively mastered and combined. The principal distinction between skills and knowledge is that the former necessarily involve performance of some activity, and are likely to have some generality, whereas the latter involve command of facts and details that are tied to specific tasks.

Organization of the Performance Domain

Three distinct principles can be used to organize and describe the relationships among work behaviors shown in Figure 1. First, as shown in Figures 1 and 2, they can be hierarchically organized in such a way that each behavior is classified as a component or an antecedent of some class of behavior at a higher level in the hierarchy. Second, as shown in Figure 3, behaviors can be classified according to sites; some behaviors occur only on the actual work site, whereas others can be assessed off-site. Third, as shown in Figure 4, behaviors can be classified as either task-related or as behaviors which do not directly relate to task performance, but which are nevertheless relevant in defining performance levels. Finally, as shown in Figure 5, behaviors can be cross-classified according to both site and task--non-task designations.

Hierarchical Organization

Work behaviors can be organized into five separate levels, ranging from a global level, which considers the impact of all behaviors at work, to a very specific level, which considers only behaviors which are directly related to levels of skill or task knowledge (Figure 2). The levels are:

1. Level I--Overall Effectiveness in Position.

2. Level II--Behavioral Component of Effective Performance. This class includes task performance, both as an individual and as a member of a team, and non-task behaviors that serve to limit effectiveness.

3. Level III--Interpersonal Relations. This class of behaviors has both a direct and an indirect (via its effect on Team Task Performance) effect on Overall Effectiveness.

4. Level IV--Job Proficiency.

5. Level V--Job-Related Skills and Knowledge.

This hierarchical arrangement has two implications. First, successful measurement at a lower level of the hierarchy does not necessarily imply that successful generalizations can be made to a higher level. Second, changes in behaviors at lower levels in the hierarchy may make themselves felt at several higher levels. Both of these principles are

straightforward extensions of the strict hierarchy of causation implied by Figure 2; levels are defined in such a way that influence always flows upward in the hierarchy.

On-site vs. Off-Site Behaviors

Behaviors at levels I, II, and III are defined only at the actual work site, whereas levels IV and V can be defined either on-site or off-site (Figure 3). This distinction is useful in two ways. First, it suggests that off-site performance measures, such as simulations, are not likely to provide adequate measures of on-site work behaviors. Second, this distinction helps to isolate the effects of contextual variables. Context is defined by the work site. As a result, contextual variables will affect only the relationships within the sub-domain of on-site behaviors. Note that relationships that are depicted as partially enclosed within the sub-domain of on-site work behaviors (e.g., the relationship between job proficiency and task performance) will also be affected by context. Thus, any arrow on Figure 1 that is partially or fully within the on-site group represents a relationship which can potentially be moderated by contextual variables.

Task vs. Non-Task Behaviors

Task behaviors can be classified as those directly involved in task performance and those that do not directly involve, or are not uniquely associated with the performance of major technical tasks (see Figure 4). Task behaviors can occur either on-site (Individual and Team Task Performance) or off-site (Skills, Knowledge, and Task Proficiency); non-task behaviors occur only at the work site. Some of these have little to do with technical tasks (Destructive/Hazardous and Down Time Behaviors), others are indirectly related to the accomplishment of some tasks (Interpersonal Relationships), and others include, but are not defined by, task performance.

The distinction between task and non-task behaviors is useful for two reasons. First, performance measurement methods that are task driven may ignore an important category of behaviors, although, as will be noted later, non-task behaviors may affect some task measures. Second, the determinants of task and non-task behaviors may be very different. In particular, traditional predictors, such as ability measures, may be ineffective in accounting for several non-task behaviors.

Cross-Classification

The three different organizing principles described above are partially redundant. For example, hierarchical levels I, III, IV, and V are each associated with only one class of task-oriented behaviors; only level II behaviors span the task - non-task classification. Site and task-based classifications are also partially redundant; only task behaviors occur both on and off site. Nevertheless, there is some utility in considering combinations of the three principles in defining unique structures for the work behavior domain.

For some purposes, it will be useful to classify work behaviors according to both their task orientation and the sites at which they occur (Figure 5). Using these two classification principles, three groups of behaviors can be defined. First, there is a set of off-site task behaviors (skills, knowledge, and job proficiency). Second, there is a set of on-site task behaviors (individual and team task performance) that may be affected by contextual variables. Third, the remaining behaviors can be classified as on-site, non-task behaviors, which can be affected by contextual variables, and which probably have quite different antecedents than task behaviors.

PERFORMANCE MEASURES AND WORK BEHAVIORS

Several types of performance measures can be identified, ranging from job knowledge tests to global evaluations provided by supervisors. These different measures will not necessarily yield equivalent information, nor will they provide equivalent coverage of the performance domain. In this section, the links between several different performance measures and the dimensions of job performance are described. Four distinct technologies are, or can be employed in measuring performance: (1) job knowledge/skills tests, (2) hands-on testing, (3) simulations, and (4) rating scales. There are several distinct applications of each technology; description of these measures is listed below.

Performance Measurement Technologies

Job Knowledge/Skills Testing

1. Paper and Pencil Tests. Most likely to be used to measure task knowledge.

2. Job Skills Tests. Performance tests can be used to assess skill components of job performance. For example, a test of skill in using power tools could be developed by having examinees complete several standard operations using designated tools. The skills measured here are more basic and general than the specific tasks that are performed in doing individual jobs. For example, electronic troubleshooting might represent a skill, whereas repairing a sonar console might represent a task.

Hands-on Testing

3. On-site, hands-on testing. Here, the subject performs one or more tasks in the normal work setting while being observed by an evaluator or evaluation team.

4. Off-site, hands-on testing. The equivalent of work sample tests, in which the subject carries out tasks using normal equipment and techniques, but in a context that is different from the work site. Van tests provide an example of this category. This type of testing measures performance of whole tasks rather than of their basic skill or knowledge components.

Simulations

5. High-fidelity simulations. Similar to off-site, hands-on testing, this technique attempts to duplicate the salient features of the work site and to obtain job proficiency measures under realistic conditions. Aircraft simulators provide an example.

6. Symbolic simulation. Also related to hands-on testing, this technique involves the use of pictorial or video materials in depicting aspects of the test or job environment. The subject's task is to indicate, using these symbolic materials, how he would carry out a task. As with hands-on tests described above, this technique measures performance in whole tasks rather than in basic task components.

Ratings

7. Task Ratings. Evaluations obtained from supervisors, peers, or the subject on how well or effectively a specific task is carried out. These judgments are distinct from the category that follows in that they are tied to individual tasks.

8. Global Ratings. Evaluations obtained from supervisors, peers, or the subject on general dimensions such as technical competence, safety, etc. These represent evaluations of the person rather than of a specific task carried out by that person.

Relating Performance Measures to Work Behaviors

There is not, for the most part, a one-to-one correspondence between specific performance measures and work behaviors outlined in Figure 1. We must therefore attempt to describe the relationships between each of the eight performance measures and the nine categories of work behavior. Initial estimates of these relationships are shown in Table 1. These relationships are characterized according to three dimensions: (1) Spurious/Casual, (2) Weak/Strong, and (3) Variable/Stable. The first dimension refers to the extent to which each of the criterion dimensions can be regarded as a direct cause of the behavior measured by the particular performance measure. The category "spurious" includes cases in which no relationship is expected; the category "casual" includes cases in which the behavior sampled by the performance measure is the same as the behavior which defines the criterion level. The second dimension refers to the expected value of the correlation between each performance measure and a hypothetical perfect measure of the work behavior. Note here that a strong relationship is not necessarily an indication of construct validity; substantial correlations may be the result of bias, shared method variance, etc. The third dimension refers to the extent to which this relationship is expected to be invariant across different contexts. Three-point scales are used for each dimension; a rating of 3, 3, 3 would indicate that the performance measure is directly caused by the criterion construct in question, and that this relationship is strong and invariant across the contexts which are likely to be encountered. The rationale for these proposed relations are presented below.

1. Paper and Pencil Tests. The bench marks which define job knowledge. Knowledge is a component of job proficiency, but individual differences in job knowledge are likely to reflect time in job and training level more than anything else, and thus will be unrelated to interpersonal and non-task behaviors, and only weakly related to task performance and overall effectiveness.

2. Job Skills Tests. The relationships proposed for job knowledge tests will also hold for this class of performance measures. One exception is in the relationship between skills and task performance. Since skills are likely to be more general and transportable than specific task knowledge, skills will be more strongly related to performance in a variety of post-training contexts than will measures of job knowledge.

3. On-site, Hands-on Testing. The bench mark that defines job proficiency. Proficiency, in turn, is in part the outcome of training, job knowledge, and job skills. There is no reason to expect a consistent relationship between this performance measure and non-test behaviors such as accidents, absenteeism, or interpersonal conflict (the same can be said for the next three performance measures). The relationship between this measure and criteria at levels I, II, and III is uncertain, since ability to accomplish tasks under controlled conditions does not always translate into effective job performance.

4. Off-site, Hands-on Testing. Since testing is removed from the usual contest, this measure will not be as strongly related to job proficiency as will on-site measures. On the other hand, if contextual cues are removed, more generalized job skills and knowledge may play a greater role. For all other criterion levels, comments made with regard to on-site testing apply here also.

5. High-fidelity Simulations. Although similar to on-site, hands-on tests, these measures deprive the subject of specific contextual cues, such as the physical layout of work, or the placement of machines and structures, and thus may lead to greater emphasis on job-related skills and knowledge. However, if sufficient fidelity can be achieved, this technique would lead to less of an emphasis on these same factors than would off-site, hands-on tests.

6. Symbolic Simulations. Since these are a step removed from actual hands-on testing, these measures may relate more strongly to training, and possibly to verbal and spatial ability, and less strongly to job proficiency than any of the hands-on or high-fidelity simulation methods.

7. Task Ratings. Judgmental measures, even when tied to specific tasks, are likely to be biased in the direction of the rater's general evaluation of the ratee (halo effect). This will vary by source (supervisor, peer, self) but is likely to lead to consistent positive correlations between ratings (both task and global) and assessments of the subject's overall effectiveness. This, in turn, will lead to a positive correlation between task ratings and assessments of the subject's overall value to the Navy. Assessments of training performance and job knowledge/skills are likely to bias task ratings, as are non-task behaviors. Finally, task ratings will be strongly related to task performance; this relationship will be stronger for ratings of individual performance than to ratings of performance as a team member, since many facets of team performance are beyond the control of the individual ratee.

8. Global Ratings. These ratings will be strongly affected by both overall task performance and non-task behaviors. Training performance will bias global ratings, particularly if training scores are known to the rater. Finally, proficiency, skills, and knowledge are all indirectly related to global ratings. Positive correlations would be expected between global ratings and any of the seven criterion levels.

Interpreting Table 1

Table 1 portrays a set of complex, multi-faceted relationships between actual performance measures and performance constructs (work behaviors). Two comments should be made regarding the interpretation of these relationships. First, for different purposes one, two, or all three dimensions might be used to characterize these relationships. Different conclusions regarding the mapping of performance measures onto the criterion space could be reached, depending on which combination of dimensions was used. Second, this set of relationships exemplifies the bandwidth-fidelity trade off. For example, all classes of work behaviors are thought to have a moderate to strong effect on global ratings. This indicates that global ratings could be used as a partial measure of any of the major classes of work behavior. On the other hand, it would be difficult to uniquely identify global ratings with any particular type or level of criterion. Hands-on tests, on the other hand, are uniquely identified with job skills and proficiency, but may not provide valid measures of other behaviors. If performance is very narrowly defined, one would hope to attain an analog of simple structure in the set of relationships depicted in Table 1. That is, under a very tight definition of performance, it would be best to develop sets of performance measures that are identified uniquely with one performance construct. In this case, the tendency of more global measures to "load" on many work behavior dimensions would represent a form of criterion contamination. On the other hand, if performance is defined very broadly, spanning large segments of the domain of work behaviors, the tendency of several measures to "load" on only one or two work behavior dimensions might be regarded as criterion insufficiency.

Measurement Implications

Table 1 implies that different performance measures will be best suited for measuring different aspects of the performance domain. For example, both paper and pencil tests and performance tests of basic skills measure job knowledge/skills and provide information which is relevant in assessing proficiency, but are only tangentially useful to higher-level criteria. Hands-on tests measure proficiency as well as knowledge and skills; off-site tasks may relate more strongly to general skills and less strongly to proficiency since they deprive the examinee of specific job-related cues provided by the work environment.

High-fidelity simulations are essentially similar to on-site, hands-on tests. They may, however, place demands on some general skills, since they also deprive the examinee of specific contextual cues that may be used in carrying out the job. Symbolic simulations may provide higher levels of fidelity in some jobs than in others. The ratings in Table 1 are based on the assumption that symbolic simulations involve significant abstraction, and thus may tap more general skills in the same way as paper and pencil tests and off-site, hands-on tests do.

Taken as a group, paper and pencil tests, job skills tests, hands-on tests, and simulations all measure the skills/knowledge/proficiency domain. Performance measures differ in the extent to which they place a premium on specific or general skills and knowledge, but do not differ in their essentials.

Task ratings are likely to be tied to task performance, although they will be influenced by rater biases and by the extent to which raters actually observe the tasks rates. Non-task factors will have their greatest impact when task outcomes are not highly visible (e.g., some inspection tasks), and when raters have limited opportunities to observe actual task performance. Under these circumstances, task ratings will reflect global evaluations of ratees. Global ratings, in turn, represent the only feasible measure of overall effectiveness. The problem with global ratings is that they are affected by task behaviors, non-task behaviors, and by a host of rater and ratee characteristics. Thus, it is difficult to know precisely what global ratings measure.

RELATING INPUT VARIABLES TO PERFORMANCE DIMENSIONS

The typical validity study relates scores on a test or a set of measures to some measure of overall performance. A better understanding of job performance can be gained by relating input variables, defined as the set of work-relevant attributes a person brings to the job, to each of the dimensions of job performance. This can be accomplished by linking the performance domain illustrated in Figure 1 to the set of attributes that are most likely to affect job performance. This linkage is illustrated in Figure 6.

Two general categories of input variables can be defined: fluid and fixed. Fluid variables are those that are likely to change as a function of a person's experiences in the Navy. Fixed variables are those that are likely to remain relatively stable over time. Fluid variables most likely to affect job performance are:

1. Training. Both formal and informal activities designed to impart job-specific knowledge and skills.

2. Experience. Includes previous work experience as well as previous billets and assignments in the Navy.

3. Motivation. Willingness to engage in, expend effort in, and take responsibility for job and task performances.

Fixed variables most likely to affect job performance are:

1. Cognitive ability. Includes general intelligence and related abilities such as verbal, mathematical, and spatial ability.

2. Psychomotor/Information-Processing ability. Includes tasks involving processing and acting on sensory signals, such as sonar operations.

3. Personality. Stable disposition to engage in certain classes of behavior, particularly those that affect interpersonal relations and risk taking.

The performance domain can be described at several different levels of specificity; dimensions can be described individually or can be grouped into homogeneous levels, sites, etc. Similarly, the relationships between input variables and performance dimensions can be described at several different levels of specificity.

Relating Input Variables to Specific Work Behavior Categories (See Table 2)

Training contributes directly to the tasks that comprise a typical job description, and has little to do with the broad domain of non-task behaviors. Since these behaviors affect task performance, measures of training success are not as effective in predicting actual task performance as they are in predicting ability to perform the task. Because of the general nature of ability, particularly cognitive ability, ability measures will be useful for predicting skills and proficiency, and will be more useful than training measures in predicting higher-level criteria which include, but are not exclusively comprised of, task performance variables.

Measures of motivation are least useful in predicting what an individual can do, and most useful for predicting what he or she will do. Individual task performance and downtime behaviors are likely to be particularly sensitive to variations in motivation; team task performance is also affected, but is beyond the complete control of the individual.

Biodata (experience) measures are most closely related to skills, but also relate to team task performance, particularly if the biodata include measures of previous performance as a team member. The moderate to strong relationship between biodata and other work behaviors is based on the track record of biodata research rather than on well-defined conceptual links between biodata items and those work behaviors. Finally, personality measures will be more useful for predicting non-task behaviors than for predicting task behaviors.

Relating Input Variables to Criterion Levels

One principle implied in Table 3 is that training and ability are more useful for predicting what a person can do than what he or she will do. Motivation, on the other hand, is useful for predicting "will do" and less useful for predicting "can do." Biodata will be useful in predicting skills and knowledge. This broad class of predictors is in fact useful for predicting all of the lower-level criteria, and as a result has considerable cumulative usefulness for predicting overall effectiveness. Personality is most useful for predicting interpersonal relations, and is markedly less useful for predicting all other lower-level criteria.

Relating Input Variables to On-Site and Off-Site Behaviors (See Table 4)

The critical distinction between on- and off-site behaviors is that the latter occurs outside of the usual physical, social, and task context. It is worth noting that off-site assessments of performance will almost always occur in weak situations characterized by high levels of motivation, and little competing task pressure. As a result, off-site behaviors will be affected by training, ability, and background/experience variables. On-site behaviors, on the other hand, reflect motivation and background; both ability and personality variables are also relevant, but will generally prove less useful as predictors.

Relating Input Variables to Task and Non-Task Behaviors (See Table 5)

Task performance most strongly reflects training, ability, and background. Motivation is important, but as is noted in the following section, this is true only for some tasks (on-site tasks). Non-task behaviors, on the other hand, most clearly reflect personality and motivational differences. Biodata will be useful for predicting some of these behaviors, particularly those related to work avoidance (e.g., absenteeism). Training and ability are marginally useful for predicting this class of behaviors.

Relating Input Variables to Off-Site Task, On-Site Task and Non-Task Behaviors (See Table 6)

Both ability and training are useful in predicting task performance, but both of these variables will be more useful for predicting performance off-site than for predicting on-site task performance. Motivation is important for predicting all types of on-site behavior, while personality is most likely to prove useful in predicting non-task behaviors. Biodata will be useful for predicting all classes of behavior, but will probably work best in predicting on-site task performance. In part, this is a function of the breadth of the biodata domain; measures of background and life experience probably relate to each of the four remaining categories of input variables.

Contextual Moderators

The relationships between input variables and performance dimensions will not be invariant, but rather will depend on the context in which work is performed. Three dimensions of work context that are likely to moderate the relationships between input and work behavior variables can be identified: (1) task responsibilities, (2) work environment, and (3) social context.

1. Task Responsibilities. Are defined as the total set of tasks performed by the worker, and by the relationship between these tasks and performance of the unit's mission. Performance in the same task could be evaluated differently, and could have different antecedents, depending on the individual's position in the unit. Every individual in the Navy is responsible for some cluster of tasks. Although individual tasks may be held in common by individuals who hold many different jobs (e.g., typing reports), they are also embedded within a position-unique task cluster that defines this contextual variable.

2. Work Environment. Is defined in terms of the physical surroundings in which work is carried out and in terms of the other activities going on at the time that the work is carried out. For example, some work environments might be hot, noisy, or cramped, while others place few demands on the worker. The availability of resources, both physical and personnel, needed to do the job would be considered an aspect of the work environment. This environment might be different during war than during peacetime, might be different at sea than at dockside, etc.

3. Social Context. Is defined by the extent to which an individual is in contact, direct or indirect, with others while performing his work. Examples of indirect contact include telephone contacts, messages left by previous shifts, etc. The social context of work will be strongly affected by the work environment and will also reflect the extent to which an individual's work is linked to the work of others. Social contacts will be maximized when the job is carried out in the presence of others, when interpersonal contacts are part of the job, or when information from others must be obtained in order to perform the job.

Effects of Context Variables

Most existing research on context effects has concentrated on situational variables that place constraints on performance, such as shortages of time, information, equipment, supplies, or support. As a result, context variables are generally thought to attenuate, or depress, the relationship between input variables and work behaviors. A more general definition of context effects is considered here. Context variables are those features of the situation in which tasks are performed that affect the relationships among input variables and work behaviors. Thus, situational constraints on performance represent only one class of context effects. Context variables might facilitate performance, or could leave performance levels unchanged, while changing the antecedents or the implications of performance. For example, in a highly cohesive work group, interpersonal skills and cognitive abilities might be less important determinants of performance than in a new work group, or in a novel work situation.

Context variables will affect several relationships. First, context will moderate the relationship between and individual's overall effectiveness and his or her general value to the Navy. A moderate level of effectiveness in a highly critical position may be more valuable than a higher level of effectiveness in another position. General value to the Navy will also depend on factors such as manpower availability, the changing mission of different naval units, and the available technology.

The focus of the framework proposed here is on the relationships among work behaviors and between input variables and work behavior. Three sets of relationships can be identified which are particularly likely to be affected by context factors. First, context is likely to moderate the relationship between job proficiency and task performance. Second, context is likely to moderate the relationship between non-task behaviors and overall effectiveness. Third, context is likely to affect the relationships between several input variables and on-site work behaviors.

Job Proficiency/Task Performance. This relationship will be strongest when individuals (1) are motivated to perform up to their capability, (2) have access to supplies, resources, and materials needed to perform tasks, and (3) are in an environment which supports the use of standard, approved procedures. Motivational research suggests that individuals perform up to their capability when they believe that equitable rewards (and punishments) are closely linked to task performances. Thus, individuals whose task performance is rarely observed, who do not think that they are being fairly evaluated, or who have not experienced link between task performance and desired outcomes, are not as likely to perform to their capacity as will other individuals working in more motivating conditions. Under motivating conditions, we might expect a low correlation between capability (job proficiency) and actual performance.

The restricted availability of resources will limit an individual's ability to translate proficiency into on-site performance. The availability of physical resources is less likely to pose a problem, since physical resources are readily inventoried and stockpiled. The availability of human resources is likely to be more problematic, since these resources are not as concrete. Nevertheless, an individual who does not receive the information, cooperation, or assistance necessary to carry out tasks may find it impossible to translate proficiency into performance.

Non-Task Behaviors/Overall Effectiveness. This relationship will be strongest when (1) the tasks performed by the individual are not critical to the mission of the unit, (2) the individual functions as a member of a highly interdependent work group, and/or (3) tight deadlines or time pressure exists.

The importance of non-task behaviors will vary inversely with the criticality of the tasks performed. All other things being equal, an individual successfully performing critical tasks will be given more leeway than another whose tasks are less critical. However, if non-task behaviors interfere with task performance, this relationship may be reversed.

Down-time behaviors will be especially critical if others' work depends on the work of the target individual, or if time pressures exist. In either case, periods of absence or of diminished functioning will be more critical than if the individual works alone on tasks which have little urgency.

Inputs/On-Site Behaviors. This is a potentially complex set of relationships. For example, the social context of work and the availability of resources will have a strong impact on the relationship between ability and on-site behaviors. When motivating conditions are present, resources are available, and group norms encourage productivity, high ability individuals are likely to perform tasks effectively, and are also likely to avoid non-task behaviors which

Personality and biodata factors may have an invariant relationship with non-task behaviors (in most situations, these input variables will be the most effective predictors of non-task behaviors), but their impact on task performance will be moderated by the social and physical context of work. For example, individual differences in adaptability to stressful work conditions may relate to both personality and previous life experiences. The impact of training and experience may also depend on working conditions. The experienced individual might adapt best to unusual working conditions, whereas the highly trained individual might perform best under conditions that most closely match those encountered in training.

The impact of motivation on overall effectiveness will be most substantial when Down-Time Behaviors contribute strongly to effectiveness, since the most common work avoidance behaviors (i.e., absenteeism, lateness) are thought to reflect low levels of work motivation. Motivation will also have a critical impact when levels of job proficiency are essentially uniform (e.g., in a homogeneous group of recent "A" school graduates); in this situation, individual differences in task performance will reflect a combination of contextual and motivational factors. Finally, motivation will have a more critical impact on jobs which are relatively easy, but which require constant attention. Vigilance tasks, such as those performed by several classes of operators (power plant, radar, etc.) provide the best example of this class of tasks.

CONCLUSIONS

1. Performance dimensions must be defined, not discovered. Inductive strategies that attempt to discover the nature of performance by observing, measuring, and analyzing performance indices are insufficient to the task, in that they involve infinite regress. Performance must be defined before it is measured; performance measures cannot be used to define performance.

2. The specification of relevant goals is a necessary first step in defining performance dimensions. Job performance is regarded as a goal-directed activity. Once goals are defined, performance dimensions can be defined in terms of work behaviors that are relevant in meeting or failing to meet those goals.

3. The dimensions of job performance in the enlisted ranks of the Navy include job proficiency, and task accomplishment dimensions similar to those that underlie many hands-on performance measures. However, the domain of job also includes interpersonal behaviors and work-avoidance behaviors that are not sampled in hands-on performance measures or their equivalents.

4. The domain of job performance can be described in terms of several hierarchically organized work behavior dimensions.

5. All existing performance measurement technologies relate to at least one of the dimensions of job performance identified herein. However, none of these measurement technologies provides complete coverage of the job performance domain. In particular, hands-on tests and their equivalents provide good measures of job proficiency and job skills, but are less useful in measuring task performance on site, or in measuring on-site, non-task behaviors that are relevant in defining job performance.

6. Input variables, such as ability, training, and personality can all be related to specific dimensions of the performance domain. None of the input variables identified here is strongly related to all performance dimensions. Rather, the likely strength of the relationships between input variables and performance dimensions varies considerably across both input and performance measures.

7. The general framework describing the relationships among input variables and performance dimensions provides a useful basis for studying the validity and adequacy of performance measures. This framework can also be used to predict the outcomes of criterion-related validity studies as the specific predictions, criteria, and job contents are varied. Finally, this framework provides a common language for relating a wide variety of projects conducted in the various Services to one another and to the broad problem of maximizing job performance in the enlisted ranks.

RECOMMENDATIONS

1. In the current report relevant goals, performance dimensions, and relationships between performance dimensions, performance measures, and input variables were defined on a purely conceptual basis. Empirical research is needed to identify in more detail relevant goals and performance dimensions. Research is also needed to test a variety of predictions made in the current report. For example, on a purely theoretical basis, it is predicted that (a) global performance ratings are strongly affected by non-task behaviors, (b) ability is more strongly related to job proficiency than to individual task

performance, (c) personality measures would prove useful in predicting non-task behaviors that affect performance. Predictions of this sort should be tested empirically.

2. No simple performance measure is likely to provide complete coverage of the performance domain. Combinations of performance measures will provide better coverage than can be obtained by any single measures. In general, a combination of ratings plus either simulations or hands-on tests will give the broadest coverage of the performance domain.

3. The use of multiple performance measures is especially critical in criterion-related validity studies. Predictors will not have the same level of validity for all performed dimensions or measures.

4. Further taxonomic research on performance dimensions is recommended. Three different taxonomic bases were applied to the set of performance dimensions defined herein (hierarchical, site, task). A variety of other approaches might be applied to yield different categorizations of performance dimensions.

5. The framework presented here is static. Research on the changes in the relationship among input and work behavior variables as an enlisted person progresses through his or her first term and beyond is recommended. For example, there is theoretical support for the hypothesis that individual differences in cognitive ability are more important determinants of performance for new recruits and apprentices than for experienced enlisted personnel. Research is needed to measure such developmental changes in the causes and nature of job performance.

Table 1

Relationships Between Performance Measures and Work Behaviors

	Performance Measure							
	Paper/ Pencil	Job Skills	On-Site Hands-on	Off-Site Hands-on	Hi-Fi Simulat.	Symbol. Simulat.	Task Rating	Global Rating
Overall Effectiveness	1,1,2	1,1,2	1,2,2	1,2,2	1,2,2	1,1,2	1,2,3	3,2,3
Task Performance	1,1,2	1,2,2	1,2,1	1,2,1	1,2,1	1,2,1	3,3,2	2,3,2
Non-Task Behaviors	1,1,3	1,1,3	1,1,3	1,1,3	1,1,3	1,1,3	1,2,2	2,3,2
Interpersonal Relations	1,1,3	1,1,3	1,1,3	1,1,3	1,1,3	1,1,3	1,2,2	2,3,2
Job Proficiency	2,2,2	2,2,2	3,3,3	3,2,3	3,3,3	2,2,2	2,2,1	2,2,2
Job Skills	3,3,3	3,3,3	3,2,2	3,3,2	3,3,2	3,3,3	2,2,2	2,2,2

Note. Rating scales are given below.

Causality

1 2 3

Spurious Causal

Strength

1 2 3

Weak Strong

Stability

1 2 3

Variable Stable

Table 2

Relating Input Variables to Work Behavior Categories

	Input Variables				
Work Behaviors	Training	Motivation	Ability	Experience	Personality
Job Skills	2	1	3	3	1
Task Knowledge	3	1	2	2	1
Job Proficiency	3	1	3	3	1
Interpersonal Relations	1	1	1	1	3
Task Performance (Individual)	2	3	2	2	1
Task Performance (Team)	2	2	2	3	2
Down-Time Behavior	1	3	1	3	3
Destructive/Hazardous Behavior	1	2	1	2	3
Overall Effectiveness	1	3	2	3	2

Note. Rating scale given below.

1	2	3
Less Useful		More Useful

Table 3

Relating Input Variables to Criterion Levels

	Input Variables				
Criterion Level	Training	Motivation	Ability	Experience	Personality
I - Overall Effectiveness	1	3	2	3	2
II - Behavioral Components	1	3	2	2	1
III - Interpersonal Relations	1	1	1	2	3
IV - Proficiency	3	1	3	2	1
V - Skills/ Knowledge	2	1	2	3	1

Table 4

Relating Input Variables to On-Site vs. Off-Site Behaviors

	Input Variables				
Behaviors	Training	Motivation	Ability	Experience	Personality
On-Site	1	3	2	3	2
Off-Site	3	1	3	3	1

Table 5

Relating Input Variables to Task vs. Non-Task Behaviors

	Input Variables				
Behaviors	Training	Motivation	Ability	Experience	Personality
Task	3	2	3	3	1
Non-Task	1	3	1	2	3

Table 6

Relating Input Variables to Off-Site, On-Site Task, and On-Site Non-Task Behaviors

	Input Variables				
Behaviors	Training	Motivation	Ability	Experience	Personality
Off-Site Task	3	1	3	2	1
On-Site Task	2	3	2	3	2
On-Site Non-Task	1	3	1	2	3

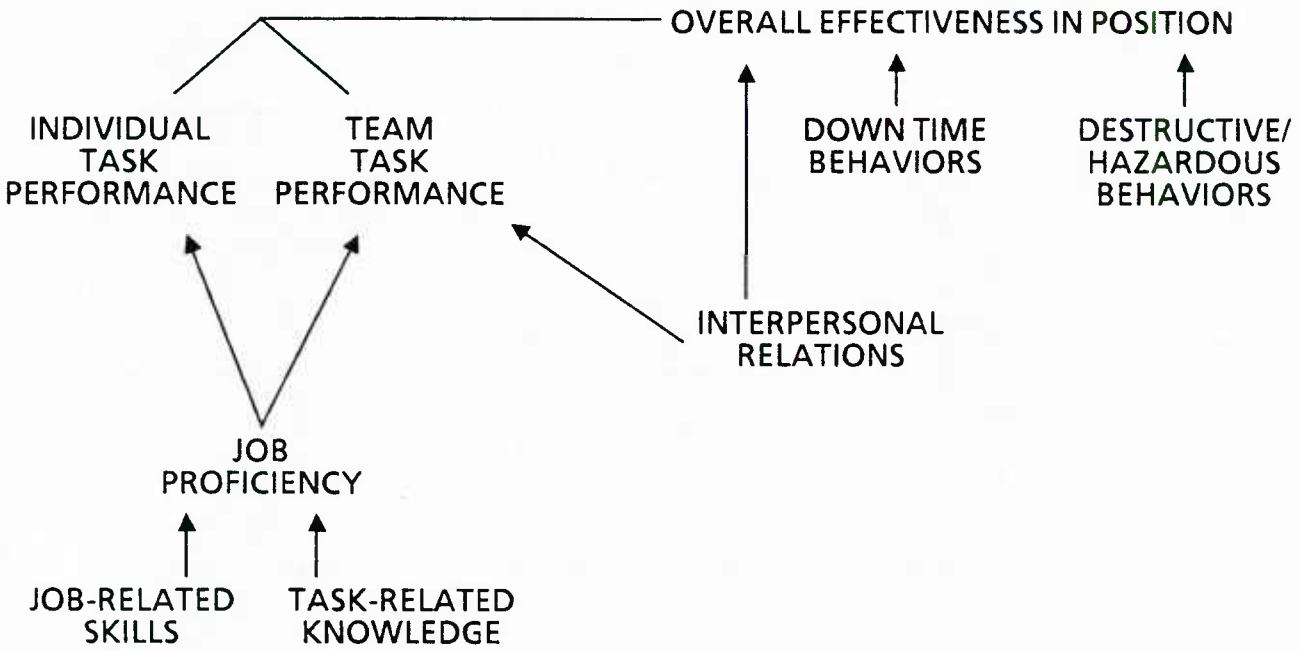

Figure 1. Dimensions of performance for Navy enlisted ranks.

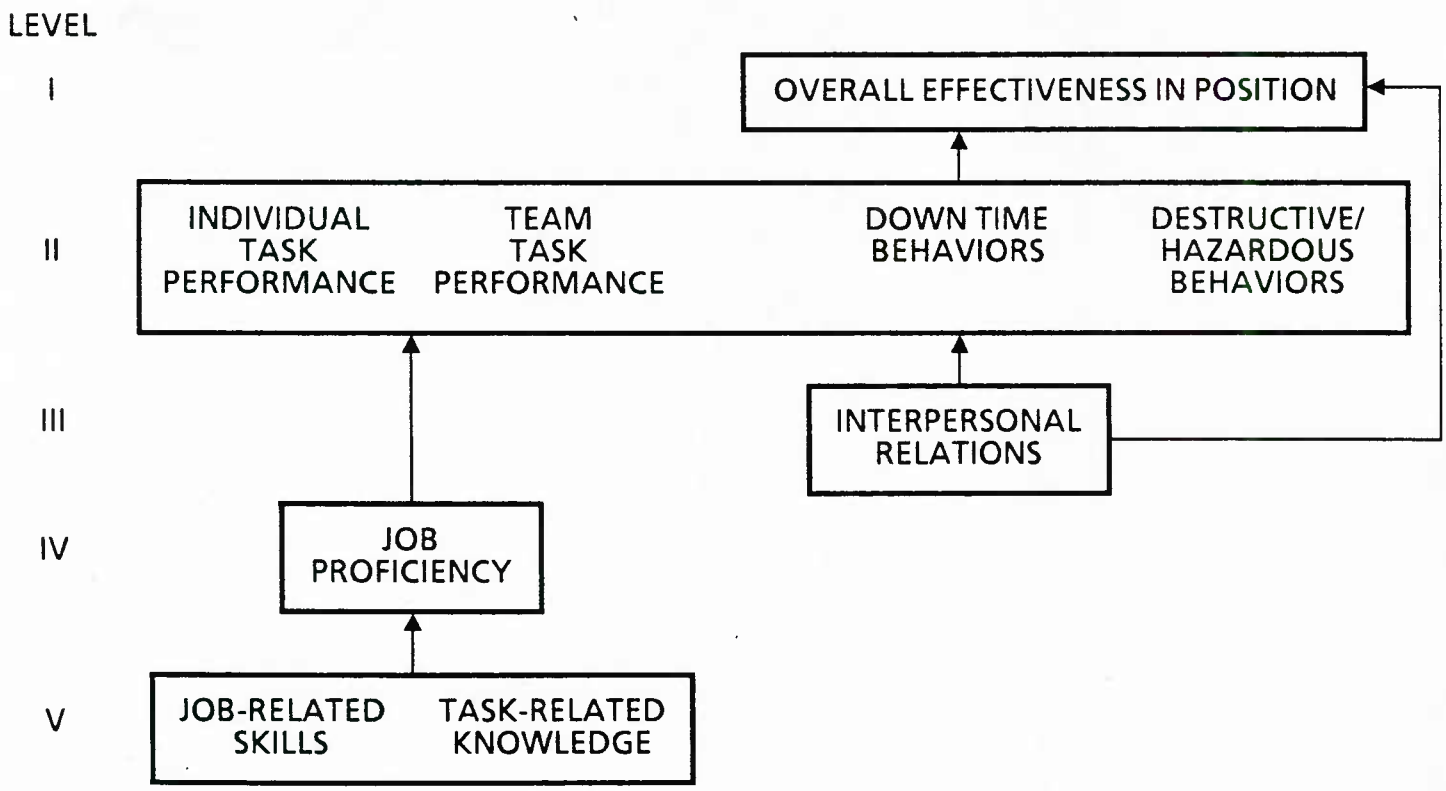

Figure 2. Hierarchical organization of work behaviors.

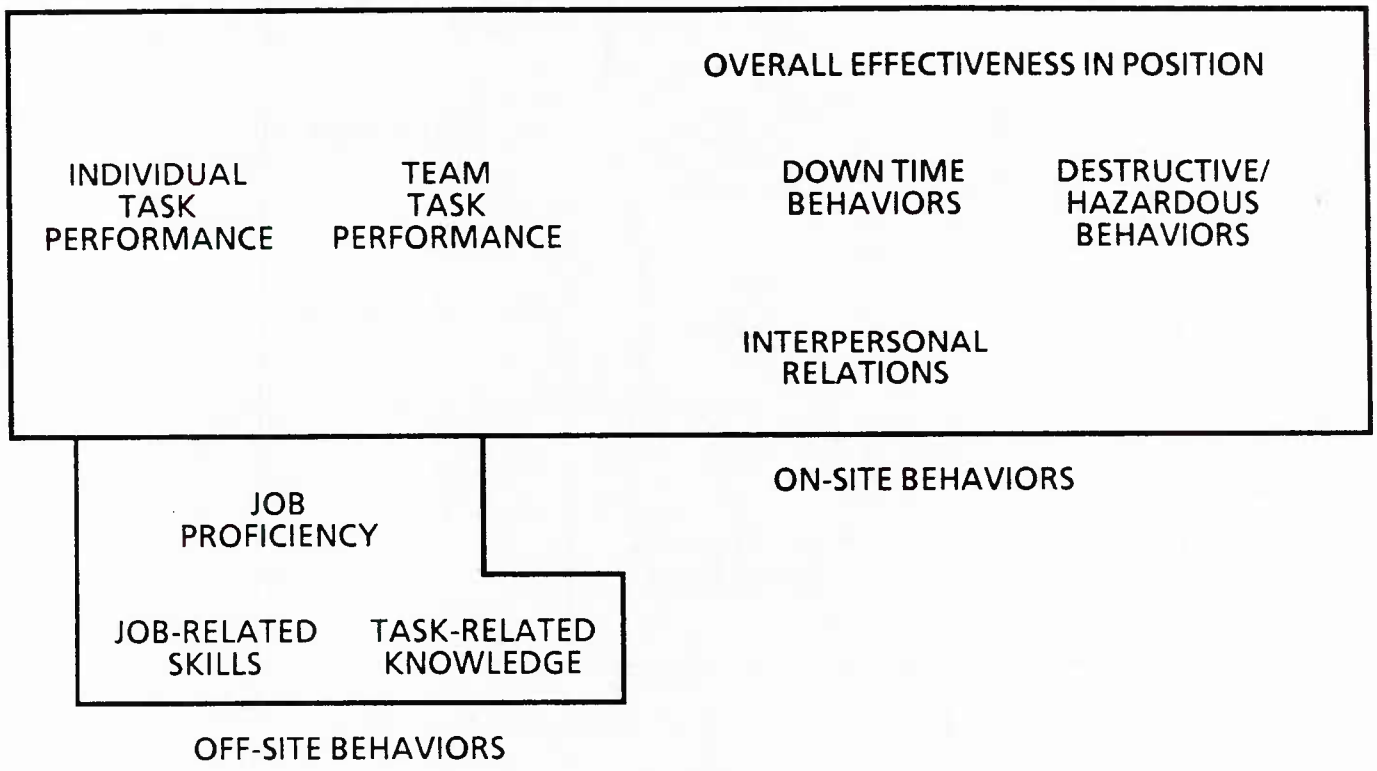

Figure 3. On-site vs. off-site behaviors.

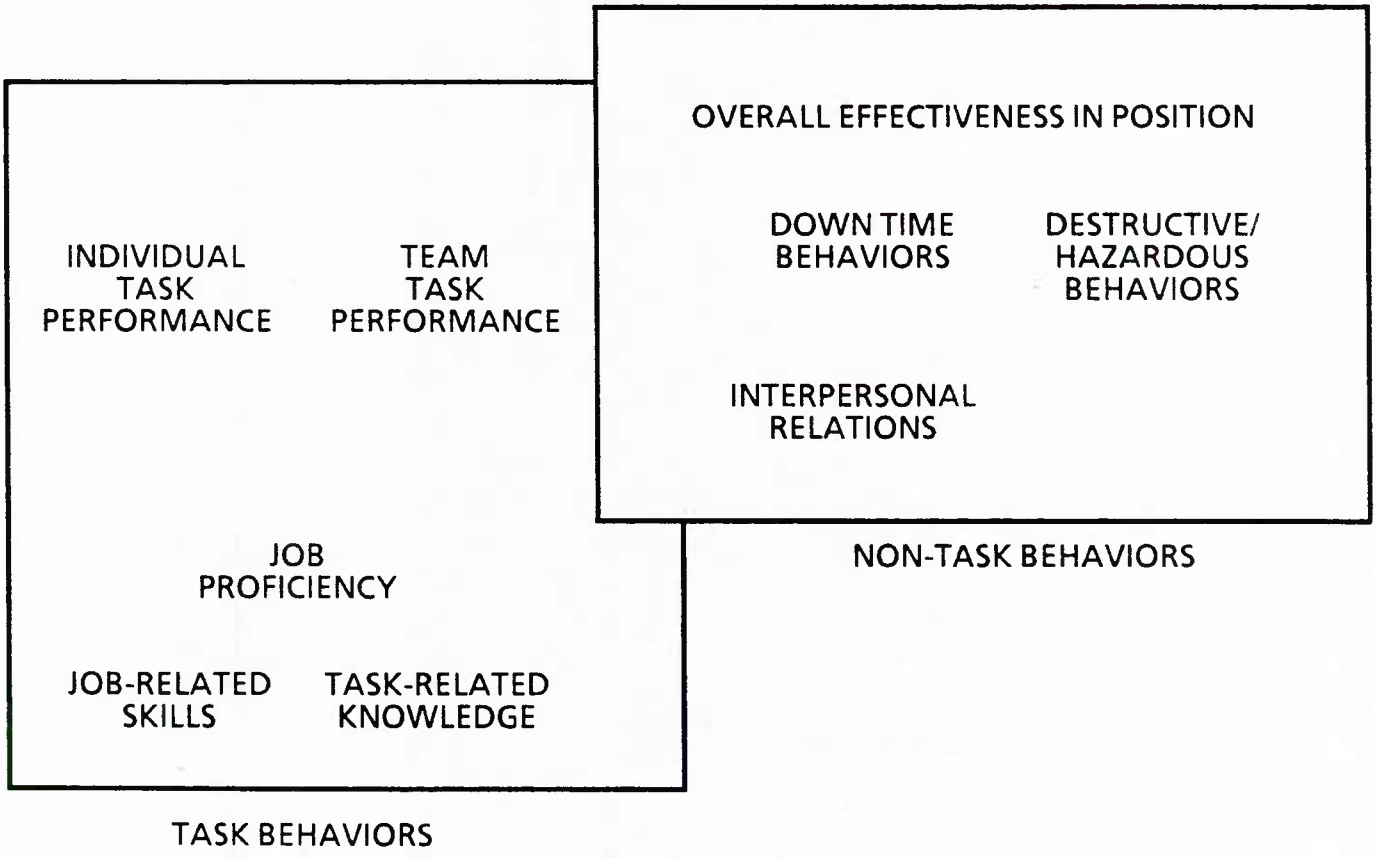

Figure 4. Task vs. non-task behaviors.

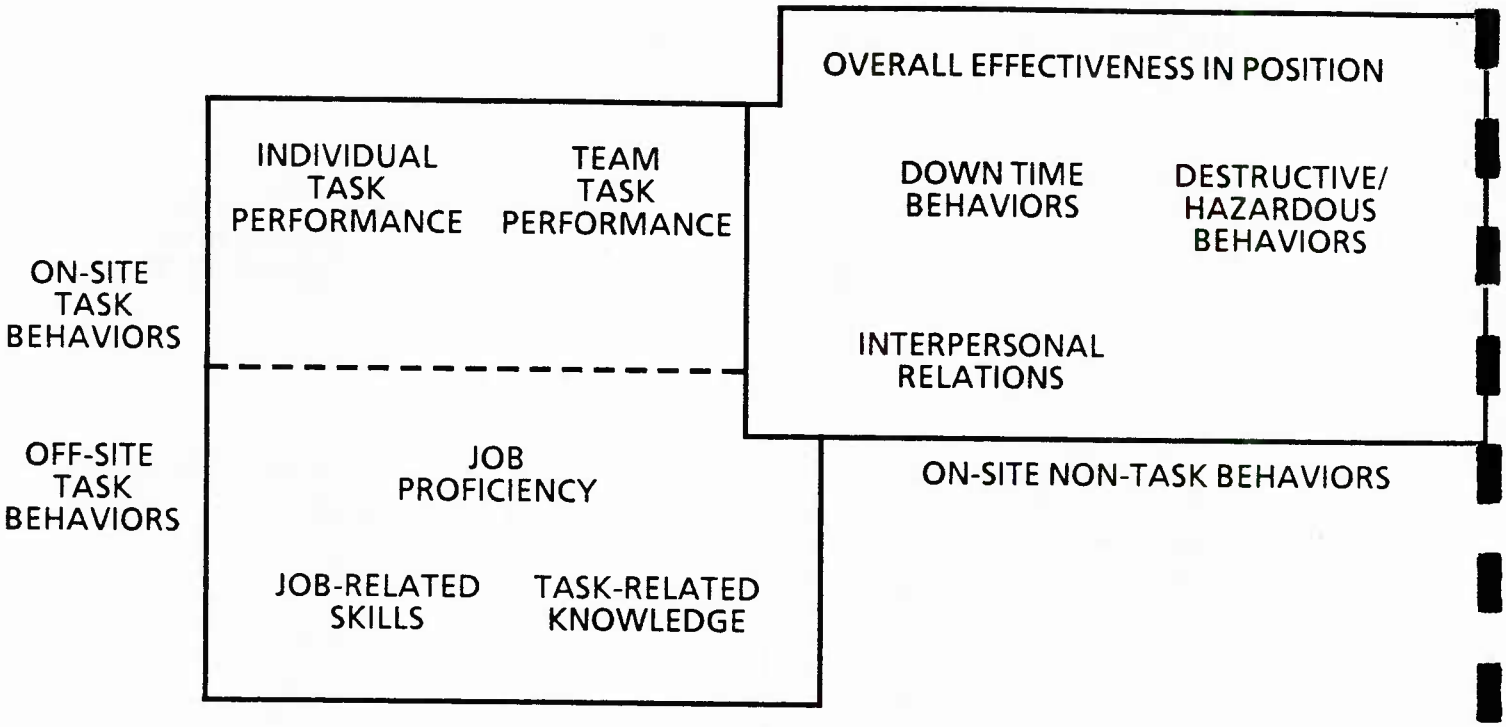

Figure 5. Cross-classification by task and site.

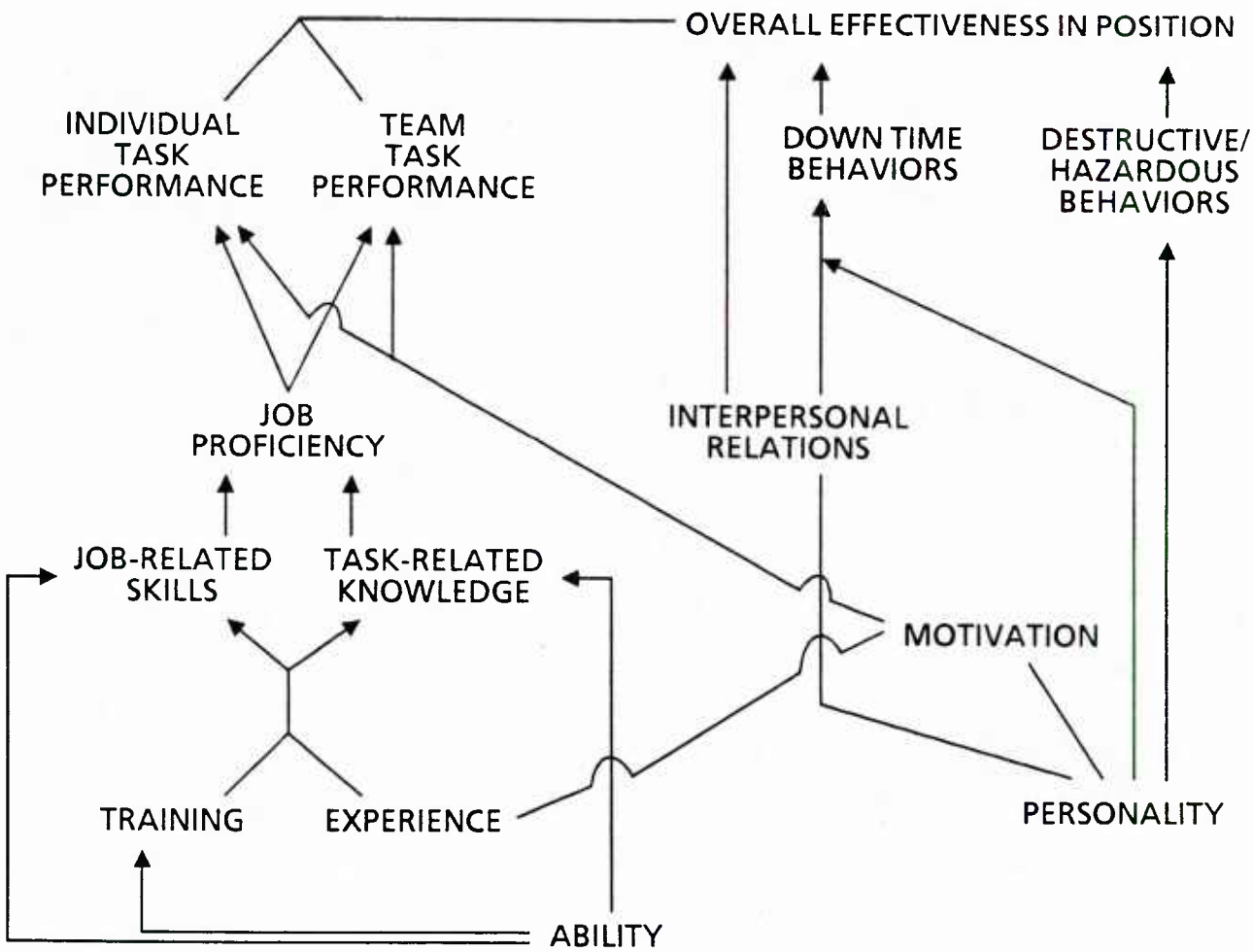

Figure 6. Linking input variables to performance dimensions.

REFERENCES

Astin, A. (1964). Criterion-centered research. Educational and Psychological Measurement, 24, 807-822.

Bass, B. (1982). Individual capability, team performance, and team productivity. In M. Dunnette & E. Fleishman (Eds.), Human performance and productivity: Human capability assessment. Hillsdale, NJ: Erlbaum.

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. Psychometrika, 45, 289-308.

Bialek, H., Zapf, D., & McGuire, W. (1977, June). Personnel turbulence and time utilization in an infantry division (HumRRO FR-WD-CA 77-11). Alexandria, VA: Human Resources Research Organization.

Campbell, J. (1983). Some possible implications of "modeling" for the conceptualization of measurement. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), Performance measurement and theory. Hillsdale, NJ: Erlbaum.

Campbell, J., Dunnette, M., Lawler, E., & Weick, K. (1970). Managerial behavior, performance and effectiveness. New York: McGraw-Hill.

Chapanis, A. (1976). Engineering psychology. In M. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally.

Christal, R. E. (1974). The United States Air Force occupational research project (AFHRL-TR-73-75). Lackland AFB, TX: USAF, AFHRL, Occupational Research Division.

Dunnette, M. (1963). A note on the criterion. Journal of Applied Psychology, 47, 251-254.

Emery, F. E., & Trist, E. L. (1960). Socio-technical systems. In C. Churchman & M. Verhurst (Eds.), Management science, models, and techniques. London: Pergammon Press.

Fleishman, E., & Quaintance, M. (1984). Taxonomies of human performance: The description of human tasks. New York: Academic Press.

Guion, R. M. (1980). On trinitarian doctrines of validity. Professional Psychology, 11, 385-398.

Guzzo, R. A., Jette, R. D., & Katzell, R. A. (1985). The effect of psychologically-based intervention programs on worker productivity. Personnel Psychology, 38, 275-293.

Hemphill, J. K. (1959). Job descriptions for executives. Harvard Business Review, 37, 55-67.

Hemphill, J. K. (1960). Dimensions of executive positions. Research monograph 98. Columbus, OH: Bureau of Business Research, Ohio State University.

Herbst, P. (1974). Socio-technical design: Strategies in multidisciplinary research. London: Tavinstock Publications.

Hunt, R. G. (1976). On the work itself: Observations concerning relations between tasks and organizational processes. In E. Miller (Ed.), Task and organization. New York: Wiley.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.

Hunter, J. E., & Schmidt, F. L. (1982). Fitting people to jobs: The impact of personnel selection on national productivity. In M. Dunnette & E. Fleishman (Eds.), Human performance and productivity: Human capability assessment. Hillsdale, NJ: Erlbaum.

James, L. (1973). Criterion models and construct validity for criteria. Psychological Bulletin, 80, 75-83.

Joreskog, K. (1969). A general approach to confirmatory factor analysis. Psychometrika, 34, 183-202.

Landy, F. J., & Farr, J. L. (1980). Performance rating. Psychological Bulletin, 87, 72-107.

Landy, F. J., & Farr, J. L. (1983). The measurement of work performances: Methods, theory and applications. New York: Academic Press.

Long, J. S. (1983). Confirmatory factor analysis: A preface to LISREL. Beverly Hills, CA: Sage.

McCormick, E. J. (1979). Job analysis: Methods and application. New York: Amacom.

McCormick, E. J., Jeanneret, P. R., & Mecham, R. C. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ). Journal of Applied Psychology, 56, 347-386.

Muckler, F. A. (1982). Evaluating productivity. In M. Dunnette & E. Fleishman (Eds.), Human performance and productivity: Human capability assessment. Hillsdale, NJ: Erlbaum.

Murphy, K. R. (1986). When your top choice turns you down: Effect of rejected offers on the utility of selection tests. Psychological Bulletin, 99, 133-138.

Nunnally, J. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. Journal of Applied Psychology, 65, 373-406.

Perrow, C. (1961). An analysis of goals in complex organizations. American Psychological Review, 26, 854-866.

Pickle, H., & Friedlander, F. (1967). Seven societal criteria of organizational success. Personnel Psychology, 20, 165-178.

Rice, A. K. (1958). Productivity and social organization. London: Tavinstock Publications.

Rice, A. K. (1963). The enterprise and its environment: A systems theory of management. London: Tavinstock Publications.

Richards, J., Taylor, C., Price, P., & Jacobsen, T. (1965). An investigation of the criterion problem for one group of medical specialists. Journal of Applied Psychology, 49, 79-90.

Ronan, W. (1963). A factor analysis of eleven job performance measures. Personnel Psychology, 16, 255-267.

Ronan, W., & Prien, E. P. (1966). Toward a criterion theory: A review and analysis of research and opinion. Greensboro, NC: The Richardson Foundation.

Rush, C. (1953). A factorial study of sales criteria. Personnel Psychology, 6, 9-24.

Salvendy, G., & Seymour, W. (1973). Prediction and development of industrial work performance. New York: John Wiley & Sons.

Schmidt, F. L., & Kaplan, L. B. (1971). Composite of multiple criteria: A review and resolution of the controversies. Personnel Psychology, 24, 419-434.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. Journal of Applied Psychology, 62, 529-540.

Schmidt, F. L., & Hunter, J. E. (1981). Employment testing: Old theories and new research findings. American Psychologist, 36, 1128-1137.

Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. (1979). The impact of valid selection procedures on work force productivity. Journal of Applied Psychology, 64, 609-626.

Seashore, S. (1975). Defining and measuring the quality of working life. In L. Davis & A. Cherns (Eds.), The quality of working life (Vol. 1). New York: The Free Press.

Smith, P. C. (1976). Behaviors, results, and organizational effectiveness: The problem of criteria. In M. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally.

Sutermeister, R. A. (1976). People and productivity. New York: McGraw-Hill.

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of goal-getting. Human Relations, 4, 3-38.

Trist, E. L., Higgins, G. W., Murray, H., & Pollack, A. B. (1963). Organizational choice. London: Tavinstock Publications.

Turner, W. (1960). Dimensions of foreman performances: A factor analysis of criterion measures. Journal of Applied Psychology, 44, 216-223.

Woodward, J. (1965). Industrial organizations: Theory and practice. London: Oxford University Press.

Woodward, J. (1970). Industrial organizations: Behavior and control. London: Oxford University Press.

DEPARTMENT OF THE NAVY
NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
(CODE___)
SAN DIEGO, CA 92152-6800

SUPERINTENDENT
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5100

0142