

AD-776 285

SELECTION OF COMPONENTS TO COMPRISE
A NAVAL PERSONNEL STATUS INDEX (NPSI)
AND A STRATEGY FOR INVESTIGATING THEIR
RELATIVE IMPORTANCE

Walter C. Borman, et al

Personnel Decisions, Incorporated

Prepared for:

Office of Naval Research

March 1974

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

Unclassified

Security Classification

AD 776 285

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)		2a. REPORT SECURITY CLASSIFICATION	
Personnel Decisions, Inc. Minneapolis, Minnesota		U	
3. REPORT TITLE		2b. GROUP	
SELECTION OF COMPONENTS TO COMPRISE A NAVAL PERSONNEL STATUS INDEX (NPSI) AND A STRATEGY FOR INVESTIGATING THEIR RELATIVE IMPORTANCE			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)			
Final Technical Report			
5. AUTHOR(S) (First name, middle initial, last name)			
Walter C. Borman Marvin D. Dunnette			
6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS	
March 1974	23 plus appendices	10	
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)		
N00014-73-C-0210	NPSI-02		
b. PROJECT NO.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)		
NR 156-020			
c.			
d.			
10. DISTRIBUTION STATEMENT			
Approved for public release; distribution unlimited.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY	
Prepared under the Navy Manpower R & D Program		Personnel & Training Research Programs Psychological Sciences Division Office of Naval Research	
13. ABSTRACT			
Rapid change in the Navy may affect the relative efficiency and effectiveness of mission or task accomplishments in Naval units and subunits. No single index is available to evaluate how such changes may affect the personnel status of Naval units. Decisions about implementing personnel programs and policies and personnel research studies could be better informed if such an index--a Naval Personnel Status Index (NPSI)--were available. (U) This report summarizes the properties that such an NPSI should have and presents the rationale for using quality indexing and policy capturing methodologies to develop an index possessing the requisite attributes. Initial steps have been taken toward such a development. A two-day workshop was held with 23 Naval officers and instructors at the Monterey Navy Postgraduate School. During this workshop, participants suggested 29 potential measures that they believed to be indications of the level of overall personnel status aboard a ship. They also judged each potential measure against scales relevant to properties deemed to be desirable for components of an NPSI and reported their perceived correlations among the 29 measures. These judgments resulted in the selection of 14 components as likely candidate measures to comprise an NPSI. Factor analysis of the perceived correlations among these 14 measures yielded five factors--I: Fitness and Readiness of Individuals; II: Discipline; III: Crew Members' Attitudes; IV: Activities During Free Time; and, V: Manning Level. The 14 components have now been used to generate a policy capturing task comprised of 100 mock ships showing various configurations on the 14 components. Naval officers will be asked to rate these mock ships according to their overall personnel status, and policy capturing methodology will be used to evaluate the components for possible use in the NPSI. (U)			

DD FORM 1473
1 NOV 65

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
(U) Personnel Status Index (U) Criterion (U) Employee Relations Index (U) Unit Effectiveness (U) Organizational Effectiveness (U) Quality of Life Measurement (U) Policy Capturing Methodology						

FINAL TECHNICAL REPORT

**SELECTION OF COMPONENTS TO COMPRISE A
NAVAL PERSONNEL STATUS INDEX (NPSI)
AND A STRATEGY FOR INVESTIGATING THEIR RELATIVE IMPORTANCE**

Walter C. Borman · Marvin D. Dunnette

March 1974

Prepared under the Navy Manpower R & D Program
of the Office of Naval Research
under Contract N00014-73-C-0210, NR 156-020.

PERSONNEL DECISIONS, INC. · MINNEAPOLIS, MINNESOTA

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

TABLE OF CONTENTS

	Page
1. BACKGROUND AND RATIONALE	
a. Results and Consensus of Conference	1
b. Results and Conclusions from Literature Review.	3
c. Assessment of Quality Indexing Methodology Properties Desired for an NPSI	4
2. THE POLICY CAPTURING METHODOLOGY	
a. General Description	7
b. Rationale for Use of Policy Capturing in NPSI Development	8
3. INITIAL NPSI DEVELOPMENT: MONTEREY WORKSHOPS	
a. Workshop Procedures	11
b. Workshop Results.	12
4. NEXT STEPS: POLICY CAPTURING JUDGMENTS ON MOCK SHIPS	
a. Generating Mock Ships	19
b. Policy Capturing Data Analysis.	19
c. Subsequent Requirements for Developing and Validating an NPSI	21
References.	24
Appendices.	25
Distribution List	

List of Tables

<u>Number</u>	<u>Page</u>
1. MEANS AND STANDARD DEVIATIONS OF CRITERION RATINGS FOR COMPONENT MEASURES	13
2. MEAN ESTIMATES OF INTERRELATIONSHIPS AMONG 29 COMPONENTS . .	15
3. MEAN ESTIMATES OF INTERRELATIONSHIPS AMONG COMPONENTS SELECTED FOR THE POLICY CAPTURING TASK	17
4. VARIMAX FACTOR LOADING MATRIX FOR THE 14 COMPONENT MEASURES SELECTED FOR THE POLICY CAPTURING TASK	18
5. INTERCORRELATIONS AMONG COMPONENTS FOR THE 100 MOCK SHIPS. .	20

1. BACKGROUND AND RATIONALE

a. Results and Consensus of Conference

In late 1972, we proposed a planning study to determine the relative feasibility of various approaches for developing a measurement system for use in determining, at any given point in time, the status of the Navy's personnel subsystem. For want of any other existing label or designation, we referred to the proposed measurement system as a Naval Personnel Status Index (NPSI). It was believed that availability of an NPSI would aid in coping with rapid changes confronting the Navy, changes which can be expected to accelerate with the advent of all volunteer forces. If developed in such a way that variations in its magnitude could yield information quickly about the effects of changes in Navy personnel policies and practices, an NPSI could provide a continuing audit or tracking of the overall personnel condition of the Navy.

A first step designed to define the problem more explicitly took the form of an informal meeting among a small number of experts representing various behavioral science disciplines and the Navy. The following persons participated in this brief two-day meeting:

Nicholas A. Bond, Jr.
Department of Psychology
Sacramento State College

Andrea Molberg
Research Assistant
Personnel Decisions, Inc.

Glenn L. Bryan
Director
Psychological Sciences Division
Office of Naval Research

Stephan J. Motowidlo
Research Assistant
Personnel Decisions, Inc.

John P. Campbell
Department of Psychology
University of Minnesota

Lee Sechrest
Department of Psychology
Northwestern University

Marvin Denicoff
Director
Information Systems Program
Office of Naval Research

Wallace Sinaiko
Research Study Director
Smithsonian Institution

Marvin D. Dunnette
President
Personnel Decisions, Inc.

Frank Smith
Director
Employee Attitude Surveys
Sears Roebuck & Company

Eric Flamholtz
Graduate School of Business
Columbia University

Gerald Thompson
Graduate School of Industrial
Administration
Carnegie Mellon University

Warren Hirsch
Department of Mathematical
Statistics
New York University

A synopsis of this meeting and a selected literature review of some possible approaches to the problem of developing the NPSI are reported by Dunnette, Milkovich, and Motowidlo (1973). That meeting touched on several issues, including such matters as: (a) whether or not the index should be reported solely as a single composite or reported as a profile made up of several relatively independent and diagnostically meaningful measures; (b) how the index should be "validated" in order to be interpretable according to evaluative connotations; (c) how "current" it would need to be to be maximally useful; (d) whether it should be used operationally or for research purposes only; (e) what units of analysis would be most desirable and feasible (e.g., tasks to be done, work groups or units, entire ships, etc.); and, finally, (f) what sampling basis might be utilized in order to evaluate properly the personnel status of the Navy or of any unit of the Navy at any given point in time. After much discussion and some exchange of post-meeting correspondence, these issues were partially "settled," and the consensus was summarized as follows by Dunnette et al (1973):

"A Naval Personnel Status Index is seen as a potentially very useful descriptive measure to help in tracking and evaluating the impact of changes in Naval personnel policies and practices as they are instituted over the months and years ahead. As a coarse indicator of change, the index should not be solely descriptive; that is, certain levels on the index should clearly be indicative of "better" levels of the personnel condition in the Navy and in particular units of the Navy than certain other levels on the index. Moreover, as a coarse indicator, the index should be capable of being expressed as a single number, but this need not preclude its being interpreted in profile form, according to its several components. In fact, its diagnostic usefulness and interpretations related to possible causes of fluctuations in its value will most assuredly be enhanced by retaining a capability for profile interpretation. Development of the basic data system necessary to comprise the index should focus on tasks or functions as primary units of analysis. Initially, many tasks should be considered, but methods should be employed to reduce their number to manageable proportion by evaluating their relative difficulty, criticality, probable frequency of occurrence; and, in particular, the relative ease of personnel intersubstitutability within them. Once such judgments have been obtained and confirmed, purposive sampling of units and functions may properly be employed to develop data gathering and reporting systems necessary for computing the index. Finally, use of the index as an operational measure to direct managerial decision making should be avoided as a means of maintaining not only the internal validity and evaluative meaning of the index, but also the integrity of information bases from which it is computed. Thus, such an index should be developed primarily for use as a research tool and not primarily for use as an operational indicator of Naval unit effectiveness. To be used

properly for personnel research and the evaluation of the products of personnel research, the index must, of course, be credible; and the maintenance of its credibility constitutes a major reason for resisting its use in an operational context." (pp. 11-12)

In sum, conferees concluded that the NPSI should:

- (a) be expressed as a single number but retain the identity of its significant component indicators for possible diagnostic purposes;
- (b) be interpretable in an evaluative sense instead of being "just" descriptive;
- (c) include components with sufficient ease of quick accessibility to insure a high degree of timeliness;
- (d) be used for research purposes only, primarily as a means of "evaluating" the relative impact of changes in personnel practices and policies or of behavioral research interventions;
- (e) focus on tasks or functions as primary units of analysis; and,
- (f) be credible and reasonably free from danger of loss of credibility through gun decking.

b. Results and Conclusions from Literature Review

Discussion of the literature surveyed by Dunnette, Milkovich, and Motowidlo (1973) was organized within four areas: (a) concepts of organizational effectiveness; (b) efforts to measure operational readiness; (c) economic value or human resource evaluation; and (d) quality composites such as "quality of life," college quality, and the like. All the methodologies reviewed are similar in that human judgments about such matters as effectiveness, readiness, or quality are in some way quantified to yield a numerical index or series of indicators.

Of the methods reviewed, we decided that the various approaches designed to develop quality composites [(d) above] seemed most likely to be practicable for developing an NPSI. A central reason for this choice of general strategies is that such approaches typically examine the judgments of experts relatively early in the developmental stages rather than seeking to impinge prior judgments or classification systems upon them. In other words, knowledgeable persons have a crucial early role in determining which components should be examined as possible quality indicators. Their judgments are used as a means of pinpointing which measures make the most sense or may be most important in contributing to an overall index such as "Quality." Quantification procedures may then be utilized at later stages in the developmental process to determine the psychometric

properties of the rationally selected components to provide clues for deriving operational measures of them and statistical procedures for combining them. In a concluding statement, Dunnette et al (1973) expressed the relative advantages of the "quality" approach for developing an NPSI as follows:

"The task of deciding what elements contribute to one's 'quality' of life seems . . . [to be] a matter about which some reasonable consensus ought to be possible. Why not, then, simply gather together various groups of Naval experts and work with them to name and to define the various indicators of good and poor personnel status? . . .

"Ways of systematically gathering such indicators might then be developed and implemented and their relative levels and variability recorded and studied over time." (p. 40)

c. Assessment of Quality Indexing Methodology Properties Desired for an NPSI

The major discrepancy between the approach being advocated above and the points of consensus derived by our conferees about the desirable properties of the proposed NPSI involves the units of analysis to be used. Our literature search has led us to be pessimistic about the feasibility of being able to synthesize successfully organizational indicators from a data base focusing on tasks or functions as the primary units. We, therefore, laid aside (for the time being, at least) our initial intention to focus on such molecular elements as tasks or functions and decided to focus instead upon a more molar unit--the ship--as the primary unit for initial study; i.e., the unit for which we hope to derive a first "experimental" NPSI.

With regard to the other desirable NPSI properties decided upon by conferees, the "quality" methodology meets them in the following ways:

- (a) Single number vs. Profile. By asking Naval experts to designate those factors they regard as important indicators of a ship's overall personnel status, we are assured that their primary focus will be on the notion of a single composite or overall index; yet, we will learn also about components that are perceived as important contributors to such an overall evaluation. Decisions about how these components relate to each other and how they may reasonably be combined into a single composite will be made by utilizing policy capturing methodology. [Details of policy capturing methodology and how we propose to use it in providing the basis for these decisions are outlined in Part II of this report.]

- (b) Descriptive vs. Evaluative. The quality index methodology obviously yields judgments based on an evaluative distinction between "good" personnel status and "poor" personnel status. Moreover, the policy capturing methodology assures that any components which contribute nothing or little to such overall evaluative judgments will be discovered and can be deleted based on results of the policy capturing data analyses. As still another way of assuring that components are selected initially which do possess definite evaluative connotations, we also have secured ratings of the relative importance to overall personnel status of each of the components suggested. These ratings are described in detail in later sections of this report.
- (c) Accessibility of Measures. As will be seen later in this report, we also have based selection of initial components on experts' ratings of the relative accessibility of measures of the components and upon ratings of their relative generality across all types of Naval vessels. (For example, we seek components that are easily accessible on all ships, not just small ones, large ones, or those unique to any other particular subset.)
- (d) Research vs. Operational Use. We have retained our intention that the NPSI will be used for research purposes only. This requirement is only relevant to the quality index methodology in that a wider range of possible components are likely to be suggested than might be the case if our expert panel were considering an NPSI which might ultimately be intended for operational uses.
- (e) Unit of Analysis. As already noted, we have altered our initial plans by moving to a molar level (i.e., "ship") as the unit of analysis instead of focusing on the molecular level of tasks or functions. [We do believe, of course, that judgments about the components entering into the overall personnel status of a ship definitely imply attention to the accomplishment on that ship of its important missions, tasks, and functions, but these estimates enter indirectly instead of directly into the formulation of the NPSI.]
- (f) Credibility and Robustness. Credibility of the NPSI and the components comprising it are, of course, assured by the very nature of the quality indexing methodology. As a further means of avoiding easily "faked" or error prone measures, our initial selection of components is also based, in part, on experts' ratings of their relative reliability and "fudgeability." These scales and results of ratings with them are described in later sections of this report.

6.

On the following pages, we outline in Section 2 the nature of policy capturing methodology and our rationale for using it in later stages of NPSI development; in Section 3, we describe initial NPSI developmental steps taken during workshop meetings held with a group of Naval officers in Monterey and results of analyzing information gained during those meetings; and, in Section 4, we describe our next step--to apply the policy capturing methodology for the purpose of inferring the importance of each component to NPSI measurement from the point of view of experienced Naval officers. In Section 4, we also recommend a post-policy capturing strategy for the final development and validation of an NPSI.

2. THE POLICY CAPTURING METHODOLOGY

a. General Description

One possibility for moving forward on the choice of component measures for a Naval Personnel Status Index (NPSI) lies in the recently developed policy capturing methodology. As we shall see, policy capturing and related techniques can provide a rational basis for weighting NPSI components to reflect experienced Naval officers' wisdom. To explain adequately how policy capturing techniques will aid in developing the NPSI, first we should briefly describe this methodology. Then we will demonstrate its applicability to NPSI development.

A number of policy capturing methods (e.g., Judgment Analysis--Christal, 1963, 1968; Madden, 1963; Naylor and Wherry, 1965) have an important objective in common--to develop models to represent the way persons weight information or cues in making global judgments about some set of stimuli. For example, when a supervisor rates the overall job performance of a number of subordinates based on his knowledge of their work in a variety of areas, policy capturing techniques might be used to discover the strategies he used in combining his knowledge of their performance in different job areas to arrive at his overall impressions of their performance. Another example: A member of a selection board charged with screening candidates for undergraduate enrollment at a college has available to her six test scores along with ratings of potential supplied by the candidates' references. The policy capturing technique can discover this person's "policy" of selecting undergraduates by pinpointing the cues she depends most heavily upon in making decisions about the candidates.

In fact, policy capturing techniques can specify a rater's selection strategy or policy any time that rater makes a single global judgment about each of a series of stimuli based on two or more quantifiable pieces of information (cues) for each stimulus. In general, it is convenient to generate a rater's policy by performing a multiple regression analysis with cues as independent variables and the rater's global judgments as the dependent variable. The vector of regression weights represents the rater's policy. Each of these weights may be thought of as an importance index. That is, a relatively high regression weight for a given cue implies that the cue is important to the rater for making judgments about dependent variable stimuli. Smaller regression weights indicate cues of lesser importance to the rater. Thus, a rater's policy or the way he uses cue information to form global impressions can be captured in the form of a multiple regression equation which relates to his use of cues in making these overall judgments.

Furthermore, the size of the multiple correlation between cues and the overall judgments indicates the degree to which the rater's evaluations are perfectly predictable by linear regression. Two factors cause this

multiple R to be less than 1.00--a configural or nonlinear rating strategy, or an inconsistent policy. Goldberg (1970) and Slovic and Lichtenstein (1971), among others, have concluded that raters are very seldom meaningfully configural in their rating strategies. That is, a linear model has accounted for almost all of the variance in judgments across a variety of rating tasks. Consequently, a vast majority of the difference between obtained R and 1.00 can be expected to be related to inconsistency in rater policy. This result can be used to assess the consistency of a rater group's composite policy. A relatively low R for a group of raters indicates that the group's composite policy is inconsistent. This in turn implies that different raters in the group are using different policies. Conversely, a relatively high R for a rater group suggests that the raters in the group have similar policies.

b. Rationale for Use of Policy Capturing in NPSI Development

We submit that policy capturing provides a very appropriate way to obtain a list of component measures potentially useful for an NPSI composite. By gathering information about what cues Naval officers weight heavily when assessing the personnel status of a ship, we can draw on the wisdom and experience of persons who are very familiar with the personnel subsystems of ships. In a sense, we are formalizing and systematizing that Naval officer wisdom when we capture the policy of officers who evaluate the personnel status of a number of ships based on a variety of information they possess about each of these ships. The way experienced Naval officers use and combine this information about ships' personnel subsystems will provide important clues about what types of measures should most legitimately be placed into an NPSI.

To obtain a clearer idea of advantages to be derived from a policy capturing analysis, let's contrast the policy capturing approach for aiding in the identification of component measures with the strategy employed by General Electric in developing the Employee Relations Index (ERI) (Merrihue and Katzell, 1955), an index similar in meaning and in purpose to the NPSI. GE researchers spent relatively little time at the conceptual tasks of selecting ERI component measures and identifying areas from which component measures should be obtained or developed. Instead, they invested considerable time insuring that the components they selected correlated well with each other and that the unit weighted composite ERI correlated "properly" with organizational outcome variables such as productivity.

Although the validation efforts just mentioned are legitimate for developing a meaningful and useful indicator of personnel status, we feel that it is a mistake to deemphasize the crucially important early component generation phase. It is very important to exhaust as completely as possible the personnel status domain by generating a wide range of ideas for components before making assumptions about what components might be more important than others.

Also, we suggest that the GE approach treated too lightly the important job of conceptually evaluating the legitimacy and the importance of these potential component measures. NPSI development should include a thoughtful evaluation of each potential component measure's conceptual suitability for inclusion in an NPSI. Furthermore, this conceptual analysis of NPSI component content should be accomplished before launching into the empirical validation of specific measures for an NPSI. We feel that the policy capturing methodology provides such a means for conceptually evaluating components for NPSI measurement. Attending to results of the policy capturing procedure should insure that final NPSI components will be measuring content areas which informed persons consider very important for determining the state of personnel subsystems.

A related reason for performing a policy capturing analysis in this project is to screen the large number of potentially important and valid NPSI components. Policy capturing results can provide an unobtrusive measure of the importance of each component (cue variable) to personnel status from the perspective of experienced Naval officers. Those components which are not weighted heavily in a policy capturing analysis can be eliminated from further consideration as part of the NPSI. By eliminating through a policy capturing procedure components which are not important indicators of personnel status in officers' minds, we will then be able to spend more time investigating measurement problems for each component seen as important for describing personnel status. And, we can forget about component measurement problems in areas seen as not so important for an NPSI to tap.

A final contribution the policy capturing methodology can make is to provide an estimate of the unanimity of opinion among officers with regard to the way they use information from components or cues to make overall judgments about personnel status. One method with this capability, Judgment Analysis (JAN), groups together raters utilizing similar policies. The grouping process results in information about the number and nature of separate policies within a rater sample. That is, from a group of raters each of whom rates the same set of stimuli and uses the same cues as other raters, JAN can identify distinct rater subgroups relatively homogeneous in terms of the weights they place on cues. Also, each subgroup's composite policy can be derived from JAN. In addition, Naylor and Schenck (1966) have suggested using the index ρ_m to assess the amount of agreement between two raters. ρ_m provides an index of interrater agreement in terms of their policies rather than their actual ratings. Either of these methods will be appropriate for identifying disagreement in "policy" among Naval officer raters.

In summary, policy capturing provides a naturalistic method for assessing the way experienced Naval officers use information pertinent to the personnel subsystem of ships to make judgments about the status of those personnel subsystems. By studying via policy capturing the way Naval officers make these judgments given the information they have available,

we can assess the importance of each piece of information used by these officers. This analysis will in turn help first to eliminate from further consideration for an NPSI, those information sources which are seen uniformly as unimportant indicators of the state of personnel subsystems on ships. And, second, policy capturing results will suggest the degree of homogeneity of officer opinions concerning the value different information sources have for measuring personnel status.

Having provided a general background statement about policy capturing and its applicability to NPSI development, we now describe more specifically the approach we favor for progressing toward the development of a useful and valid NPSI. In the course of this discussion, we report results of a pilot investigation workshop held with Naval officers. The purpose of the workshop was to generate potential components for an NPSI composite.

3. INITIAL NPSI DEVELOPMENT: MONTEREY WORKSHOPS

a. Workshop Procedures

An obvious first step toward selecting components for an NPSI is to generate a list of potentially useful component measures. To accomplish this, two PDI staff professionals organized and led a two-day workshop designed to generate such a list. Approximately 23 Naval officers attending the Navy Postgraduate School in Monterey, California participated in the workshop sessions. (See Appendix 1 for a copy of the workshop booklet containing a participant roster and the schedule of workshop events.) In addition, a small group of instructors at the Postgraduate School attended some or all of the sessions. The breadth and amount of experience possessed by this group seemed to be adequate for our purposes. For example, the workshop group contained officers with shipboard experience in small patrol vessels, destroyers, cruisers, and carriers.

Furthermore, although no officer of captain rank or above attended the workshops, we felt that the group's experience level was high in terms of knowledge about a wide variety of potential NPSI component measures. Most officers were quite familiar with such data systems as NEC and 3M. Also, most participants were knowledgeable about the configuration of a variety of Bureau of Naval Personnel data potentially valuable for an NPSI composite. Finally, all officers knew a considerable amount about data routinely kept aboard individual ships; and, they had a good feel for possible ways to combine existing data to yield useful component measures.

During the introduction period, PDI staff members discussed NPSI project progress and outlined the upcoming workshop activities. Participants were then divided into two subgroups to generate ideas for potential component measures. Officers in one of the subgroups were encouraged to recall ships they had been assigned to or were familiar with and then to think about the status of that ship's personnel subsystem. Then, these officers were asked to record the cues, information, or factors which led them to assess the personnel status of that ship the way they did. This method resulted in the generation of 20-30 components potentially applicable to NPSI measurement. The PDI leader of the other subgroup requested that his officers generate potential NPSI component measures directly. Thus, officers in this subgroup introduced component measure ideas directly to others in the subgroup. Then, each suggestion was discussed and eventually accepted, rejected, or refined by other group members. This procedure also resulted in 20-30 ideas for NPSI components.

The two PDI staff members then pooled the ideas they had received from their groups. There was considerable overlap between the two groups' output in terms of the kinds of components suggested and even the actual measures proposed for each component. The pooling operation yielded 29 measures or areas¹ seen as important for NPSI measurement.

¹See Appendix 2 for the NPSI component list. Where possible, the PDI staff members specified the way a component was meant to be measured. It was impossible to know exactly how some of the components might be measured, however. In these cases, we stated as precisely as possible the content or nature of the component.

These 29 components were named and listed in preparation for the next day's rating session. The second day of the workshop was devoted to two rating tasks. First, officer participants were asked to evaluate each component on five criteria of importance and utility for use in an NPSI composite. (See Appendix 3 for an example of the rating task.) The five criteria were: Importance; Reliability; Accessibility; Generalizability; and "Fudgeability." The purpose of obtaining these criterion ratings was to assess officers' perceptions of the "goodness" of each component. These opinions of component quality on a number of criteria enabled us to identify several components which showed promise for contributing to an NPSI composite.

The second rating task required each officer to respond to one of two formats designed to estimate the intercorrelations among component measures. That is, instead of obtaining the empirical relationships among components by actually obtaining measures from a large number of ships and then computing their intercorrelations, we asked officers to provide us with estimates of these intercorrelations. One of the formats required raters to picture a ship which stood extremely favorably along a particular component (1 of the 29). The rater was then asked to rate from 1 (extremely unfavorable) to 9 (extremely favorable) the probable standing of that ship along the other 28 components. The other format was identical except that the rater was asked to picture a ship which stood extremely unfavorably on a particular component. (See Appendix 3 for the two rating protocols.)

This rating task provided a relatively uncomplicated method for obtaining each rater's estimate of the relationships among components. Furthermore, we received these estimates from a number of officers independently and from two separate rating formats.

b. Workshop Results

- i. Ratings. Results of the criterion ratings appear in Table 1. It is of interest to note that the general importance level of the components was seen as quite high. Only six components have mean importance ratings of below 3.00 on a scale of 1 to 5. Also, the generalizability results were encouraging. All means on this criterion were near 2.00--fully generalizable. Thus, according to these Naval officers, there should be few problems obtaining component scores across different ship types.

We used these criterion results to select from the total list of 29 those components which seemed most promising for providing high quality NPSI measurement. These 14 component measures are now in the process of being studied in the policy capturing phase of our research. Here are the specific criteria used in selecting the 14 components for inclusion in the policy capturing rating task.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF CRITERION RATINGS
FOR COMPONENT MEASURES (N = 20)

	C R I T E R I A									
	Importance		Reliability		Generalizability		Accessibility		Fudgeability	
	M	SD	M	SD	M	SD	M	SD	M	SD
1. Keenlist	4.05	.74	3.75	.54	1.95	.22	4.00	.00	3.50	.59
2. Mast	3.25	.83	3.20	1.03	1.85	.36	3.80	.40	2.80	1.03
3. R/NR	3.05	1.07	3.80	.51	1.90	.30	3.40	.49	3.50	.59
4. Q/NQ	4.10	.83	2.95	.80	1.85	.36	2.70	.64	2.50	1.07
5. UAR	3.85	1.01	3.35	.79	1.95	.22	3.70	.46	2.70	.84
6. Reenlist-FT	4.20	.87	3.70	.64	1.95	.22	3.95	.22	3.60	.58
7. Turnover	3.90	.83	3.55	.74	2.00	.00	2.90	.54	2.90	.89
8. Court	2.80	1.08	3.55	.86	1.95	.22	3.75	.43	3.25	1.04
9. Manning(NEC)	3.90	.99	3.65	.57	1.95	.22	3.50	.59	3.45	.50
10. Pass Rate	3.45	.86	3.75	.54	2.00	.00	3.60	.49	3.60	.58
11. Rec/Eligible	3.15	1.06	3.25	.83	1.95	.22	3.05	.38	2.40	1.11
12. Courses	3.30	1.05	3.10	.77	1.95	.22	2.75	.54	1.80	.98
13. Group	3.25	1.09	2.25	.89	2.00	.00	2.30	.64	1.50	.50
14. Sick Rate	2.95	.86	3.10	.94	1.80	.40	3.16	.49	2.10	.94
15. Sack Time	1.74	.96	1.35	.57	1.74	.44	1.70	.71	1.16	.36
16. Morale	3.00	1.38	1.95	.67	1.90	.30	2.20	.60	1.35	.73
17. Pers Capability	3.20	1.33	2.35	.57	1.95	.22	2.45	.59	1.50	.87
18. Leadership	4.15	.79	2.35	.48	2.00	.00	2.20	.40	1.75	.70
19. Food	3.35	.96	2.60	.58	1.95	.22	2.65	.79	2.00	.77
20. Drugs	2.95	1.24	3.15	.96	1.95	.22	3.75	.54	2.75	.94
21. Admat	2.90	1.18	2.95	.86	1.95	.22	3.95	.22	3.05	.86
22. Discharge	2.80	1.12	3.60	.73	2.00	.00	3.65	.48	3.40	.66
23. Grievance	3.30	1.19	2.80	1.08	1.95	.22	2.55	.97	2.80	1.03
24. Maintenance	3.50	1.12	2.90	.93	1.85	.36	3.65	.48	2.05	.74
25. Trained	3.10	1.14	2.30	.78	2.00	.00	2.80	.60	1.60	.80
26. Transfer	3.35	1.24	2.95	.80	2.00	.00	2.90	.62	2.15	1.11
27. Sat-Pa	3.60	.80	2.30	.64	1.95	.22	2.15	.36	1.90	.62
28. Sat-SM	3.60	.80	2.40	.66	2.00	.00	2.15	.36	2.00	.71
29. Enlist-Fit	3.45	1.50	3.40	.73	2.00	.00	2.85	.57	3.25	.62

- (a) Importance. For the most part, components rated as relatively important were selected for the policy capturing rating task. However, two components (Sack Time and Discharge) were included as "dummy variables." That is, these two variables were seen as not very important by officers in the workshop. If the policy capturing analyses suggest that they are very important in terms of affecting ratings of personnel status, we will have reason to question the policy capturing results. Thus, these two components have been included, in a sense, to check on the policy capturing procedure.
 - (b) Generalizability. Components were required to be potentially generalizable across ship type.
 - (c) Reliability, Accessibility, and "Fudgeability." We attempted to select components which showed up relatively favorably in these three areas (except for the two "dummy variables").
 - (d) Range of content. Components were selected such that the total group of 14 represented a wide range of content (e.g., discipline, crew satisfaction) and diverse ways of measuring components (e.g., objective measures, ratings by crew members).
- ii. Perceived relationships among components. Table 2 displays the results of rating task #2 designed to estimate the correlations² among components. We also obtained a crude estimate of inter-rater agreement with respect to the independent rater estimates of the relationships among components. Three intraclass correlation coefficients (R_1) (Haggard, 1958) were computed--one each for ratings provided on the "favorable" rating task, the "unfavorable" rating task, and the two combined.³ The results: R_1 (favorable) = .44; R_1 (unfavorable) = .53; R_1 (combined) = .63. The R_1 s provide

²These "correlations" were obtained by first applying an arbitrary linear transformation to the ratings. For the task in which raters were to imagine a ship with an extremely favorable standing along a component and then to estimate the same ship's probable standing on other components, a rating of 9 (extremely favorable) was transformed to +1.00. The other transformations were 8 = +.75; 7 = +.50; 6 = +.25; 5 = 0; 4 = -.25; 3 = -.50; 2 = -.75; 1 = -1.00. For the task in which raters were to imagine a ship with an extremely unfavorable standing along a component and then to estimate the same ship's probable standing along the other components, the transformations were, of course, reversed. That is, 1 = +1.00; 2 = +.75, etc. Then, all estimates of correlations among components were averaged to obtain the coefficients appearing in Table 2. We recognize that this is, at best, an extremely crude approach to approximating the correlations among these variables. Nonetheless, it does give a first look at how experienced Naval officers perceive these variables to covary. As such, the matrix provides a better basis for concocting the hypothetical ships in the policy capturing phase than would be provided by our own a priori judgments.

³For the "combined" analysis, ratings from the "unfavorable" task were first reflected. That is, 1s were transformed to 9s, 2s to 8s, etc.

TABLE 2
MEAN ESTIMATES OF INTERRELATIONSHIPS AMONG 29 COMPONENTS*

Components	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1. Reenlist																													
2. Mast	13**																												
3. R/NR	42	20																											
4. Q/NQ	47	22	81																										
5. UAR	35	85	27	27																									
6. Reenlist-FL	77	17	27	28	07																								
7. Turnover	59	25	32	32	45	52																							
8. Court	13	57	20	15	77	00	17																						
9. Manning (NEC)	42	13	52	57	06	36	17	06																					
10. Pass Rate	50	06	47	60	15	32	25	06	45																				
11. Rec/Eligible	45	32	47	50	25	55	30	14	02	-25																			
12. Courses	46	17	25	35	13	33	06	15	08	63	71																		
13. Group	44	23	10	10	27	44	42	19	08	10	21	36																	
14. Sick Rate	23	35	02	15	46	29	21	25	08	08	27	29	29																
15. Sack Time	13	10	08	13	27	23	04	08	08	13	29	50	27	59															
16. Morale	69	42	21	44	46	69	29	27	29	27	48	41	61	61	52														
17. Pers Capability	42	29	52	67	29	48	31	19	54	38	66	39	11	38	41	64													
18. Leadership	63	40	15	25	54	65	25	27	15	21	57	43	38	39	36	71	41												
19. Food	33	02	06	10	06	33	13	02	08	04	04	05	13	38	05	52	18	25											
20. Drugs	38	46	42	25	52	42	25	50	15	17	21	27	30	57	36	66	41	45	06										
21. Admat	22	16	28	47	19	19	25	19	41	25	18	27	16	20	22	36	41	50	23	20									
22. Discharge	25	47	58	33	56	25	03	50	17	11	18	14	20	32	15	34	27	32	09	59	00								
23. Grievance	25	14	33	28	17	22	17	11	19	11	34	14	18	36	13	39	20	43	25	23	09	28							
24. Maintenance	13	-06	22	44	09	06	28	00	42	22	25	22	15	30	22	30	50	32	00	13	28	06	09						
25. Trained	08	-14	03	25	-08	06	03	-06	03	06	07	25	07	09	07	18	11	45	00	14	47	-06	-03	-13					
26. Transfer	42	22	19	22	25	39	22	14	28	19	59	02	43	32	22	61	23	59	43	36	03	16	38	03	-06				
27. Sat-Pa	69	25	11	33	36	61	25	08	28	33	61	30	52	48	27	64	25	70	48	52	09	13	50	09	03	81			
28. Sat-SM	67	25	19	33	28	58	17	08	25	28	32	16	55	36	27	48	34	48	27	43	16	16	09	16	-06	56	75		
29. Enlist-Fit	31	17	28	50	25	22	06	06	31	50	36	41	05	25	15	14	50	05	02	27	25	16	19	28	-03	06	09	25	

*See Footnote 2, page 14, for an explanation of how these correlations were computed.

**Note: Decimal points omitted.

a rough idea of the variation in the estimates of interrelationships among components provided by different raters for the same component pairs, compared to the overall variation in the estimates across all raters and all component pairs. The relatively high positive R_1 s suggest that officer raters agreed substantially among themselves about the relative magnitudes of correlations among these components.

Although the arbitrary linear transformations discussed in Footnote 2 make untenable a serious consideration of absolute correlation size, it is of interest to note that many of the estimated relationships among components are near zero or only moderately positive. This pattern suggests (provided the officer raters are right) that it may be unrealistic to expect an NPSI composite to consist of a number of highly related components, all measuring the same narrow unidimensional construct. An NPSI may need to be multidimensional in nature to sample adequately the multifaceted domain legitimate for an NPSI to tap. The complexity of a composite NPSI will increase the problems in validating it, but the complex multidimensionality of such a composite should also increase the richness of the information provided by the index.

- iii. Factor analysis. As an extremely tentative step toward exploring the possible dimensionality of the measurable personnel status domain, we factor analyzed the correlation matrix generated by the estimates of the relationships among components provided by officers in rating task #2. For this analysis, we used only those 14 components to be included in the policy capturing phase. Table 3 depicts the 14 X 14 correlation matrix representing officers' estimates of the relationships among those 14 components. We present the factor analysis results in Table 4. To obtain these results, a principal components factor analysis first was performed. Then, the first five factors were rotated using the varimax criterion. These rotated factors appear in Table 4. Although we must keep in mind the source of the original correlations on which this analysis is based, the results do suggest very meaningful clusters of components.

Factors

- I - Fitness and Readiness of Individuals
- II - Discipline
- III - Crew Member Attitude Toward Officers and the Navy
- IV - Free Time Activities
- V - Manning Level

Again, it must be emphasized that these factor analysis results are based on data from a correlation matrix generated in a very unusual manner. Yet, the results do suggest that it may be meaningful and useful to represent the personnel status construct in terms of multiple dimensions of the type appearing in the above factor analysis results.

TABLE 3

MEAN ESTIMATES OF INTERRELATIONSHIPS
 AMONG COMPONENTS SELECTED FOR THE POLICY CAPTURING TASK*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Reenlist														
2. Mast	13**													
C 3. UAR	35	85												
O 4. Reenlist-FT	77	17	07											
M 5. Turnover	59	25	45	52										
P 6. Manning (NEC)	42	13	05	36	17									
O 7. Pass Rate	50	05	15	32	25	45								
N 8. Courses	46	17	13	33	06	08	63							
E 9. Sack Time	13	10	27	23	04	08	13	50						
N 10. Leadership	63	40	54	65	25	15	21	43	36					
T 11. Discharge	25	47	56	25	03	17	11	14	15	32				
S 12. Maintenance	13	-06	09	06	28	42	22	22	22	32	06			
13. Sat-Pa	69	25	36	61	25	28	33	30	27	70	13	09		
14. Enlist-Fit	31	17	25	22	06	31	50	41	15	05	16	28	09	

*See Footnote 2, page 14, for an explanation of how these correlations were computed.

**Decimal points omitted.

TABLE 4

VARIMAX FACTOR LOADING MATRIX FOR THE 14 COMPONENT MEASURES
SELECTED FOR THE POLICY CAPTURING TASK

		F A C T O R S					Communality
	I	II	III	IV	V		
1. Reenlist	-35*	12	87	-01	08	90	
2. Mast	-04	90	12	-01	-05	83	
3. UAR	-02	94	21	-11	10	94	
4. Reenlist-FT	19	03	87	-08	03	81	
5. Turnover	02	23	59	26	39	63	
6. Manning (NEC)	-44	03	28	17	56	62	
7. Pass Rate	-82	-01	30	-08	10	78	
8. Courses	-60	04	24	-64	-08	83	
9. Sack Time	-05	12	10	-83	15	74	
10. Leadership	07	35	69	-48	13	84	
11. Discharge	-15	70	08	-08	-01	53	
12. Maintenance	-12	-02	02	-27	91	91	
13. Sat-Pa	-06	14	79	-26	-02	71	
14. Enlist-Fit	-80	20	-03	-06	17	71	

*Decimal points omitted.

4. NEXT STEPS: POLICY CAPTURING JUDGMENTS ON MOCK SHIPS

a. Generating Mock Ships

We will ask approximately 50 Naval officers to participate in the policy capturing rating project. The procedure will require each officer to study individual mock ships, each ship characterized by scores on the 14 components identified earlier. Raters will be instructed to evaluate each "ship's" personnel subsystem in terms of its overall status. Each rater will provide these overall personnel status ratings for 100 mock ships. Appendix 4 contains the complete policy capturing task.

The 1400 component scores were generated by a computer program developed by Rosse (1970) for the CDC 6600. All scores are whole numbers ranging from 1 to 9. Each of the 14 component variables has a distribution close to that used in the stanine system (across 100 ships, mean = 5.0, standard deviation = 2.0). Also, the correlation matrix representing the relationships among components for the policy capturing task is approximately equivalent to the correlation matrix generated by officer raters in rating task #2 described earlier. That is, the correlations among components for these mock ships is very similar to the estimates of "real" component intercorrelations provided by officer raters during the Monterey workshop meetings. Table 5 presents the correlation matrix associated with the component scores in the policy capturing task. Notice that the matrices in Tables 3 and 5 are quite similar. Thus, we have succeeded in making the rating task "realistic" in the sense that actual relationships among components for our mock ships map very closely the perceived relationships reported by the experienced officers who participated in the Monterey workshop sessions.

b. Policy Capturing Data Analysis

After the 50 officers have performed the policy capturing ratings, we will analyze the resulting data with the following two major objectives in mind:

- (a) to derive inferences about the importance of individual components in terms of the way each influences the overall personnel status ratings of mock ships.
- (b) to assess the homogeneity of the policies generated by Naval officers.

In general, the data analysis will involve computing multiple regression equations for each officer in the rater group. Each equation will represent that rater's policy related to assessing the personnel status of ships. Then, we will use JAN and perhaps the pm statistic to assess the similarity in raters' policies. Other analyses will be performed as necessary to develop inferences about the nature of the officer groups' composite policy, and about the importance of each component in Naval officers' minds for determining the state of the personnel subsystem aboard ships.

TABLE 5
INTERCORRELATIONS AMONG COMPONENTS FOR THE 100 MOCK SHIPS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Reenlist														
2. Mast	20*													
C 3. UAR	28	90												
O 4. Reenlist-FT	83	06	13											
M 5. Turnover	54	26	41	50										
P 6. Manning (NEC)	40	06	04	35	13									
O 7. Pass Rate	58	08	13	34	23	42								
N 8. Courses	47	12	20	30	04	03	62							
E 9. Sack Time	11	17	27	20	02	09	03	46						
N 10. Leadership	61	43	53	53	23	12	22	47	40					
T 11. Discharge	27	55	52	27	02	12	06	16	12	38				
S 12. Maintenance	15	-04	08	03	23	42	22	19	23	26	03			
13. Sat-Pa	69	29	33	62	21	35	35	39	26	75	13	05		
14. Enlist-Fit	32	20	24	10	14	40	64	47	10	02	22	24	05	

*Decimal points omitted.

After the policy capturing analyses have been completed, PDI's staff will select from the list of 14 those components which look most promising for further study. One of the main criteria for selecting components at this stage will be the magnitude of the importance weights attached to component measures by the policy capturing procedure. Thus, the policy capturing analyses will result in a final list of components seen as possessing maximum potential for providing meaningful information within the NPSI framework. Those components surviving the policy capturing screening should definitely be of significant importance in experienced officers' minds for indicating the standing of a ship's personnel subsystem. In fact, by accepting for the NPSI those components which carry weight with experienced Naval officers in the way they assess ships' personnel status, we insure that the areas covered by our NPSI are of practical and operational significance. The next step is to specify more exactly how each component is to be measured. Below, we discuss that step and others focused on the future development and validation of an NPSI.

c. Subsequent Requirements for Developing and Validating an NPSI

For purposes of explaining our post-policy capturing approach to NPSI development, it will be helpful to distinguish among three different steps in the total NPSI development process.

- 1) selecting components;
- 2) norming component measures;
- 3) validating the index.

First, policy capturing results should suggest the most likely components for an NPSI. Components selected for this NPSI will be those which are most consistently influential in affecting personnel status evaluations in the policy capturing rating task; i.e., those components which are inferred to be most important for NPSI measurement based on policy capturing results. Policy capturing procedures and the rationale for using this technique have been outlined previously.

The next step⁴ in NPSI development will involve norming or scaling each component measure included in the NPSI. To understand how norming relates to policy capturing in our NPSI development strategy, we must first outline two assumptions made in the policy capturing phase. The first assumption is that ships can be graded fairly (from 1 = extremely poor to 7 = extremely good) on each component measure appearing in the rating format.

⁴It may be necessary first to perform varying degrees of instrument development work on components selected for the final NPSI. For example, if the component SAT-PA is chosen, we will need to develop a questionnaire to tap that component. On the other hand, means for measuring components such as Enlist-Fit are already available.

That is, we assume that some reasonable method can be devised for assigning scale scores to raw component scores to represent fairly the standing of ships along individual components. A related assumption is that officers can eventually agree in their assignments of scale scores (1 - 7) across the complete range of raw scores possible on each component. For example, we assume all officers would agree that a certain unauthorized absence rate (UAR) for a ship should be described as a "4" (average) in terms of the level of UAR represented by that raw score.

We feel that the first of these two assumptions is justified. We will outline a general norming strategy which should be effective in scaling raw component scores such that they reflect properly the state of the personnel subsystem along each component. Making the second assumption is necessary to keep separate the policy capturing and the norming activities. We see this separation as extremely desirable. It seems reasonable to assume that experienced officers can agree about scale score assignments to raw scores once they have information about the specific way in which component measures are to be taken and knowledge of the distribution of raw scores for each component across a representative sample of actual Navy ships. However, officer raters will not have this kind of information during the policy capturing task. Therefore, to introduce raw scores as component score cues in the policy capturing task forces raters to make scaling decisions for which they do not possess good information. It appears more reasonable to provide component scale scores in the policy capturing task and, thus, to assume that officers can agree on raw scale score conversions once they possess more knowledge of the actual component measures and of the raw score distributions across Navy ships. Furthermore, including scale scores rather than raw scores in the policy capturing rating task insures that the policy capturing procedure can accomplish what it was designed to accomplish. Different raters are able to make decisions about the personnel status of mock ships based on identical knowledge of component scores for those ships. Thus, inferences about the importance of components for NPSI measurement can be made knowing that raters had available the same cues. And, the scaling or norming of raw component scores, an important activity in its own right, does not confound the policy capturing results. Instead, norming is dealt with separately.

The norming process requires that experienced Naval officers inspect the distribution of raw scores for each component measure across a representative sample of real ships. These officers may then convert the component raw scores to perceived scale scores. For example, in dealing with the component Unauthorized Absence Rate (UAR), officers would assign each of the seven scale scores to a range of raw scores on the basis of information provided by the raw score distribution and their beliefs about the relative seriousness of different levels of UAR. We anticipate that officers will not be able to make scale assignments strictly on the basis of the distribution of raw scores. That is, the scale score "7" will not be assigned necessarily to the range of raw scores representing the top

two percent of the distribution of scales, the scale score "6" to the range of raw scores representing the next highest eight percent (normal distribution percentages) in the raw score distribution, etc. Instead, officers will need to use their shipboard experience in addition to information provided by the raw score distributions to establish reasonable norms. Completion of this norming process will result in a standardized system for obtaining NPSI scores for ships. Next, norm tables can be developed to transform raw component scores into scale scores. Then, for any given ship, the scale scores can be added together to provide an overall NPSI score for that ship.

The third and final stage in NPSI development involves the validation of the NPSI composite.⁵ We must demonstrate that scores on the index are reasonably reliable and that they are related (in the proper direction) to events, activities, policy changes, etc., which should affect a ship's personnel status. It is also important to show that the NPSI is not related significantly to random events which should have nothing to do with a ship's standing on the index. Essentially, we are describing the construct validation process. Although the criteria are very stringent for comprehensively validating an instrument within the construct validation framework, this strategy seems most appropriate for assessing the quality of information provided by the NPSI. Thus, we recommend performing a "for research only" empirical validation of the NPSI following a construct validation strategy.

⁵Our factor analysis results (p. 18) suggest that it might be reasonable to represent the personnel status domain with two or more subscales, each covering a different portion of this domain. If multiple NPSI subscales are established, the same validation principles discussed in this section will apply to each subscale.

REFERENCES

- Christal, R. E. JAN: A technique for analyzing individual and group judgment. Lackland Air Force Base, Texas: 6570th Personnel Research Laboratory, Aerospace Medical Division, February, 1963. (PRL-TDR-63-3, ASTIA Document AD-403 813)
- Christal, R. E. Selecting a harem--and other applications of the policy capturing model. Journal of Experimental Education, 1968, 36, 35-41.
- Dunnette, M. D., Milkovich, G. T., and Motowidlo, S. J. Possible approaches for development of a Naval Personnel Status Index (NPSI). Interim Technical Report, Contract N00014-73-C-0210, NR 156-020. Minneapolis, Minnesota: Personnel Decisions, Inc., 1973.
- Goldberg, L. R. Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. Psychological Bulletin, 1970, 73, 422-433.
- Haggard, E. A. Intraclass correlation and the analysis of variance. New York: Dryden Press, 1958.
- Madden, J. M. An application to job evaluation of a policy capturing model for analyzing individual and group judgments. Lackland Air Force Base, Texas: 6570th Personnel Research Laboratory, Aerospace Medical Division, May, 1963. (PRL-TDR-63-15)
- Merrihue, W. V., and Katzell, R. A. ERI--Yardstick of employee relations. Harvard Business Review, November-December 1955, 33, 91-99.
- Naylor, J. C., and Schenck, E. A. pm as an "Error-free index of interrater agreement." Educational and Psychological Measurement, 1966, 26, 815-824.
- Naylor, J. C., and Wherry, R. J. The use of simulated stimuli and the "JAN" technique to capture and cluster the policies of raters. Educational and Psychological Measurement, 1965, 25, 969-986.
- Slovic, P., and Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

A P P E N D I C E S

1. Participant Workshop Booklet
for NPSI Development Sessions
2. Possible Indicators of Personnel Status
3. Rating Tasks for Workshop Sessions
4. Policy Capturing Rating Task

A P P E N D I X I

AGENDA FOR PERSONNEL STATUS INDEX DEVELOPMENT SESSIONS

Thursday, 17 January, 1974

<u>Time</u>	<u>Topic</u>	<u>Discussant</u>
1500	Total Group: Introduction of panel participants and Personnel Decisions' staff members	Dunnette & Borman
1530-1630	Total Group: General Introduction to the personnel status index (PSI) project including a description of progress to date	Dunnette & Borman
1630-1830	Two Separate Groups: "Brainstorming" sessions to explore present and potential measures likely to be useful for the PSI Group A: Group B:	Dunnette Borman
1830-1900	Total Group: Discussion of separate group accomplishments and debriefing of panel participants	Dunnette & Borman
1900-1930	Cocktails	
1930	Dinner	

Friday, 18 January, 1974

0800-0830	Total Group: Review of progress and introduction to the rating task	Dunnette & Borman
0830-1100 (approximately)	Total Group: Perform the rating task (working individually)	Borman

January 17 - 18, 1974

WORKSHOP PARTICIPANTS
R o s t e r

Allen, Galen B.	Commander
Bedow, Robert J.	Lieutenant Commander
Bell, Merlin G.	Lieutenant Commander
Benediktsson, Phillip W.	Commander
Browne, Peter A.	Lieutenant Commander
Brunelli, John F.	Lieutenant
Daeschner, William E.	Lieutenant Commander
Emerson, George A.	Lieutenant
Glover, William F. H. III	Lieutenant Commander
Hamilton, Gerald K.	Lieutenant Commander
Larson, Jerold J.	Commander
May, Douglas E.	Lieutenant Commander
Mozingo, Thomas P.	Lieutenant
Nemmers, Robert S.	Lieutenant Commander
Olson, Stephen R.	Lieutenant
Paddock, John S.	Lieutenant
Quinlan, John H.	Lieutenant
Rantschler, Robert D.	Lieutenant Commander
Schraeder, John Y.	Lieutenant Commander
Stephan, Robert A.	Commander
Swan, James W.	Lieutenant Commander
Tate, Thomas	Commander
Wylie, Walter J.	Lieutenant Commander

January 17 - 18, 1974

A BRIEF SYNOPSIS OF THE NAVAL PERSONNEL STATUS PROJECT

In late 1972, we proposed a planning study to determine the relative feasibility of various approaches for developing a measurement system for use in determining, at any given point in time, the status of the Navy's personnel subsystem. For want of any other existing label or designation, we referred to the proposed measurement system as a Naval Personnel Status Index (NPSI). It was believed that availability of an NPSI would aid in coping with rapid changes confronting the Navy, changes which can be expected to accelerate with the advent of all volunteer forces. If developed in such a way that variations in its magnitude could yield information quickly about the effects of changes in Navy personnel policies and practices, an NPSI could provide a continuing audit or tracking of the overall personnel condition of the Navy. Obviously, such an index was not currently "lying on the shelf" ready for immediate application. Its development would need to be undertaken with sophistication, capitalizing on the best available wisdom from Naval and behavioral science resources.

A first step designed to capitalize on such wisdom was an informal meeting of a small group of experts representing the behavioral sciences and the Navy. We anticipated that the participants in this meeting would define the problem more explicitly, comment on the relevance of their existing research to its solution, and point us toward other investigators pursuing research along similar dimensions.

The meeting occurred on 16 and 17 November 1972 at the Washingtonian Inn, Gaithersburg, Maryland. The participants included the following persons:

Bond, Nicholas A., Jr.	Department of Psychology Sacramento State College
Bryan, Glenn L.	Director, Psychological Sciences Division Office of Naval Research
Campbell, John P.	Department of Psychology University of Minnesota
Denicoff, Marvin	Director, Information Systems Program Office of Naval Research
Dunnette, Marvin D.	President Personnel Decisions, Inc.
Flamholtz, Eric	Graduate School of Business Columbia University
Hirsch, Warren	Department of Mathematical Statistics New York University
Molberg, Andrea	Research Assistant Personnel Decisions, Inc.

Motowidlo, Stephan J.	Research Assistant Personnel Decisions, Inc.
Sechrest, Lee	Department of Psychology Northwestern University
Sinaiko, Wallace	Research Study Director Smithsonian Institution
Smith, Frank	Director, Employee Attitude Surveys Sears Roebuck & Company
Thompson, Gerald	Graduate School of Industrial Administration Carnegie Mellon University

Prior to the meeting, Bryan and Dunnette had each speculated about the so-called "ideal" properties that a Naval Personnel Status Index might possess. Dunnette, for example, had written that such an index:

'might be a single index, or it may be several (such as economic indicators), or it may be a composite. It should be expressed as a common number (such as 0 to 100, or, perhaps with 100 set as a baseline) and should reflect . . . a kind of combination of operational readiness, overall quality, efficiency, cost, and current level of 'success' of the personnel subsystem of any given Naval organizational unit. [Some measures currently used but poorly investigated are reenlistment rates, total number of billet (job) vacancies, and various attitude and opinion indicators from periodic Navy-wide questionnaire surveys. If just a single unit, such as a destroyer, is considered, its total range of possible missions can be listed and ordered according to criticality and probability of occurrence. Naturally, carrying out most such missions doesn't require all personnel aboard the craft, but some few--usually of high criticality--do. How 'lean' can the personnel subsystem of a destroyer be and still 'get by' when the occasional crisis occurs?]

"Such an index should be both robust and sensitive Its magnitude should not be subject to incidental or random events . . . nor dependent upon complicated parametric assumptions. On the other hand, it should be capable of reflecting important . . changes in Naval policies and practices. Information making up an NPSI ought to be easily and readily available so that a continuing audit (probably on a sampling basis) can be maintained over time. However, it should not be "fakeable" or, if fakeable, the methodology must somehow be such as to avoid inducing faking. . . . Certainly, low values on the NPSI should not be viewed as leading to punitive actions, though the head of an organizational unit should view low values 'with concern' and, in time, have capability for upgrading them."

The results of the Galthersburg meeting and a brief review of relevant literature are included in the Interim Technical Report [Possible Approaches for Development of a Naval Personnel Status Index (NPSI)] which each of you has received. Of the many approaches reviewed in that report, each demands human judgment. No approach or index, no matter how objective appearing nor how complete its degree of quantification, is free from human and potentially subjective judgment during some stage of its development.

One promising approach is characterized by methods which have yielded summary estimates of the so-called quality of life in different states. We wish, during these brief sessions with you, to investigate the feasibility of using such an approach in our Naval Personnel Status Index Project. We are asking the participants in these meetings to share with us the wisdom of their Naval experiences and opinions about what factors seem to be particularly important, relevant, and meaningful as indicators of good or poor personnel status in the Navy. We wish to use the information you give us to help in deciding which measures may be feasible and available from Naval units on a systematic and continuing basis. Those that are deemed to be capable of such measurement will then be introduced on an experimental or pilot basis in a number of Naval units, and values of the index will be compared with other information available from those same units.

January 17 - 18, 1974

A CHECK LIST OF SOME POSSIBLE INDICATORS OF PERSONNEL STATUS

accidents and injuries
amount of agreement among
members of a unit
amount of downtime
amount of effort
amount of work output

conflict among members of a unit
costs of keeping unit operational
decisiveness of a unit
difficulty of job assignment
disciplinary actions

disturbances in the unit
efficiency of the unit
external stress for a unit
feedback in a unit
following formal chain of command

formal evaluation of the unit
gripping and grievances
importance of job assignment
interchangeability of tasks
among unit members
job performance ratings

knowing what's required to
get assignments done
leadership in a unit
level of training
liking for other people in
the unit
measures of life satisfaction

measures of satisfaction
measures of unit morale
military bearing and appearance
number of fully qualified
personnel in unit
openness in a unit

participation rates in various
programs
pride of people for their unit
problem solving and trouble-
shooting skills
quality of work output
quickness of response to
assignments

rate of turnover in the unit
reenlistment rate
reputation of a unit
scores on qualifying tests
sickness

stability of the unit
unauthorized absences from unit
unit performance ratings
use of drugs
willingness to do whatever is
necessary

January 17 - 18, 1974

**CRITERIA FOR
ASSESSING THE "GOODNESS" OF COMPONENT MEASURES**

Importance: Importance attached to the measure for determining a ship's personnel status.

Reliability: Degree to which the component measure is free from distortion or error (either constant error or variable error).

Availability: Degree to which the component measure is available or accessible to researchers.

Generalizability: Ability of the component measure to be applicable to a variety of ship types.

"Fudgeability": Degree to which a measure of this component can be fudged or "gun-decked" by a ship's personnel.

A P P E N D I X 2

POSSIBLE INDICATORS OF PERSONNEL STATUS

1. (REENLIST) Reenlistment rate
2. (MAST) Non-judicial punishment rate
3. (R/NR) Ratio of rated to non-rated personnel
4. (Q/NQ) Ratio of qualified to non-qualified personnel
5. (UAR) Unauthorized absence rate
6. (REENLIST-FT) First tour reenlistment rate
7. (TURNOVER) Long-term stability of personnel
8. (COURT) Court martial rate
9. (MANNING NEC) Manning level ratio--describing the proportion of billets manned by qualified persons according to the NEC data system
10. (PASS RATE) Percent of persons taking rating exams who pass
11. (REC/ELIGIBLE) Percent of persons who are eligible by longevity who are recommended for rating
12. (COURSES) Rate of correspondence course participation
13. (GROUP) Rate of participation in group activities
14. (SICK RATE) Number of times per man per ship per month sick call reports
15. (SACK TIME) Percent time spent sleeping while off duty
16. (MORALE) Rating of crew morale by top officers
17. (PERS CAPABILITY) Overall rating of personnel capability by top officers
18. (LEADERSHIP) Perceived leadership effectiveness by crew members
19. (FOOD) Food and personal services evaluation
20. (DRUGS) Reported drug usage
21. (ADMAT) Overall Admat inspection grade
22. (DISCHARGE) Discharges other than honorable (including reference to special BuPers Code)
23. (GRIEVANCE) Grievances directed to others outside of command (e.g., Congress)
24. (MAINTENANCE) Percentage of maintenance actions deferred due to insufficient manning or expertise (from 3M data system)
25. (TRAINED) Percent of required training completed by officers
26. (TRANSFER) Requests for transfer, man/unit time
27. (SAT-PA) Satisfaction with present assignment (as measured by confidential questionnaire to sample of officers and enlisted)
28. (SAT-SM) Satisfaction with shipmates (as measured by confidential questionnaire to sample of officers and enlisted)
29. (ENLIST-FIT) GCT + ARI + Educational Level + Mean Semi-Annual Evaluation (all scores standardized) for all or for a sample of enlisted persons averaged across the unit

-35.

A P P E N D I X 3

The rating task in this packet requires you to make judgments about the "goodness" of the component measures we have identified during this conference. We ask you to make your ratings along a number of criteria relevant to judging the potential goodness of the components.

Importance

Importance you attach to this measure for determining a ship's personnel status.

5 = very important; crucial indicator

4 = quite important

3 = important but other component measures may be more important

2 = of some importance, but other component measures are definitely more important

1 = of little importance

	<u>Component</u>	<u>Importance</u>
1.	_____	_____
2.	_____	_____
3.	_____	_____
4.	_____	_____
5.	_____	_____
6.	_____	_____
7.	_____	_____
8.	_____	_____
9.	_____	_____
10.	_____	_____
11.	_____	_____
12.	_____	_____
13.	_____	_____
14.	_____	_____
15.	_____	_____
16.	_____	_____
17.	_____	_____
18.	_____	_____
19.	_____	_____
20.	_____	_____
21.	_____	_____
22.	_____	_____
23.	_____	_____
24.	_____	_____
25.	_____	_____

Reliability

Degree to which the component measure is free from distortion or error (either constant error or variable error).

- 4 = extremely error free; very reliable measure
- 3 = might be subject to some error but should be a relatively pure measure
- 2 = definitely contains some error, but the distortion is not great enough to cause a serious lack of reliability
- 1 = contains so much error that scores on the component measure are seriously distorted

Component

Reliability

1.	_____	_____
2.	_____	_____
3.	_____	_____
4.	_____	_____
5.	_____	_____
6.	_____	_____
7.	_____	_____
8.	_____	_____
9.	_____	_____
10.	_____	_____
11.	_____	_____
12.	_____	_____
13.	_____	_____
14.	_____	_____
15.	_____	_____
16.	_____	_____
17.	_____	_____
18.	_____	_____
19.	_____	_____
20.	_____	_____
21.	_____	_____
22.	_____	_____
23.	_____	_____
24.	_____	_____
25.	_____	_____

Generalizability

Ability of the component measure to be applicable to a variety of ship types.

2 = completely generalizable; pertains to all ship types

1 - component measure would not be generalizable to all ship types (state the ship types for which the component measure would apply)

Component

Generalizability

1.	_____	_____
2.	_____	_____
3.	_____	_____
4.	_____	_____
5.	_____	_____
6.	_____	_____
7.	_____	_____
8.	_____	_____
9.	_____	_____
10.	_____	_____
11.	_____	_____
12.	_____	_____
13.	_____	_____
14.	_____	_____
15.	_____	_____
16.	_____	_____
17.	_____	_____
18.	_____	_____
19.	_____	_____
20.	_____	_____
21.	_____	_____
22.	_____	_____
23.	_____	_____
24.	_____	_____
25.	_____	_____

Availability

Degree to which the component measure is available or accessible to researchers.

- 4 = measure of this component is available now--routinely gathered information
- 3 = easy to get or develop such a component measure
- 2 = with considerable effort, one could develop a measure for this component
- 1 = very difficult or impossible to develop a measure for this component

Component

Availability

1.	_____	_____
2.	_____	_____
3.	_____	_____
4.	_____	_____
5.	_____	_____
6.	_____	_____
7.	_____	_____
8.	_____	_____
9.	_____	_____
10.	_____	_____
11.	_____	_____
12.	_____	_____
13.	_____	_____
14.	_____	_____
15.	_____	_____
16.	_____	_____
17.	_____	_____
18.	_____	_____
19.	_____	_____
20.	_____	_____
21.	_____	_____
22.	_____	_____
23.	_____	_____
24.	_____	_____
25.	_____	_____

"Fudgeability"

Degree to which a measure of this component could be fudged or 'gun-decked' by a ship's personnel.

4 = impossible to fudge

3 = difficult to fudge

2 = could be fudged with some prior knowledge that the measure was to be taken

1 = very easily fudged

Component

"Fudgeability"

1.	_____	_____
2.	_____	_____
3.	_____	_____
4.	_____	_____
5.	_____	_____
6.	_____	_____
7.	_____	_____
8.	_____	_____
9.	_____	_____
10.	_____	_____
11.	_____	_____
12.	_____	_____
13.	_____	_____
14.	_____	_____
15.	_____	_____
16.	_____	_____
17.	_____	_____
18.	_____	_____
19.	_____	_____
20.	_____	_____
21.	_____	_____
22.	_____	_____
23.	_____	_____
24.	_____	_____
25.	_____	_____

"Favorable" Set

With the rating procedure below we are attempting to determine the degree to which components might be correlated or might tend to go together in the real world; i.e., for a ship. In performing these ratings, we ask you to consider a ship which possesses a very high or favorable "score" on a particular component. (We will specify the components you have been assigned in a moment.) For example, if the component is "number of AWOLs per person on the ship per year," imagine a ship which stands extremely favorably along this component. With this hypothetical ship in mind, we would like you to rate each other component in terms of the likelihood that the ship would show up favorably on those components. Please use the following code:

- 9 } Certain or almost certain the ship would stand favorably on this
8 } component; based on my experience, standing on these two compo-
7 } nents is definitely correlated positively.
- 6 } Standing along this component should have little if anything to
5 } do with a ship's very high standing on the other component;
4 } based on my experience, the two components are weakly correlated,
if at all.
- 3 } Certain or almost certain the ship would stand unfavorably on this
2 } component; based on my experience, standing on these two compo-
1 } nents is definitely correlated negatively.

blank: Don't know or not familiar with one or both components being considered.

* * * * *

Use the three-point range within each category to express finer distinctions. For example, a 7 rating would mean that a ship's standing along that component might tend to be favorable, but that it is by no means a certainty.

"Unfavorable" Set

With the rating procedure below we are attempting to determine the degree to which components might be correlated or might tend to go together in the real world; i.e., for a ship. In performing these ratings, we ask you to consider a ship which possesses a very low or unfavorable "score" on a particular component. (We will specify the components you have been assigned in a moment.) For example, if the component is "number of AWOLs per person on the ship per year," imagine a ship which stands extremely unfavorably along this component. With this hypothetical ship in mind, we would like you to rate each other component in terms of the likelihood that the ship would show up favorably on those components. Please use the following code:

- 9 } Certain or almost certain the ship would stand favorably on this
8 } component; based on my experience, standing on these two components
7 } is definitely correlated negatively.
- 6 } Standing along this component should have little if anything to do
5 } with a ship's very low standing on the other component; based on
4 } my experience, the two components are weakly correlated, if at all.
- 3 } Certain or almost certain the ship would stand unfavorably on this
2 } component; based on my experience, standing on these two components
1 } is definitely correlated positively.

blank: Don't know or not familiar with one or both components being considered.

* * * * *

Use the three-point range within each category to express finer distinctions. For example, a 7 rating would mean that a ship's standing along that component might tend to be favorable, but that it is by no means a certainty.

Total Component List

Components being "rated"

		—	—	—	—
1.	_____	—	—	—	—
2.	_____	—	—	—	—
3.	_____	—	—	—	—
4.	_____	—	—	—	—
5.	_____	—	—	—	—
6.	_____	—	—	—	—
7.	_____	—	—	—	—
8.	_____	—	—	—	—
9.	_____	—	—	—	—
10.	_____	—	—	—	—
11.	_____	—	—	—	—
12.	_____	—	—	—	—
13.	_____	—	—	—	—
14.	_____	—	—	—	—
15.	_____	—	—	—	—
16.	_____	—	—	—	—
17.	_____	—	—	—	—
18.	_____	—	—	—	—
19.	_____	—	—	—	—
20.	_____	—	—	—	—
21.	_____	—	—	—	—
22.	_____	—	—	—	—
23.	_____	—	—	—	—
24.	_____	—	—	—	—
25.	_____	—	—	—	—

A P P E N D I X 4

INSTRUCTIONS FOR THE RATING TASK

Today we are asking you essentially to evaluate the state of the personnel subsystem for each of 100 ships. Each "ship" will be a hypothetical one, described by scale scores on 14 components thought to be related to the state of the personnel subsystem or to the personnel status of that ship.

As you will see in a moment, your task is to study the 14 component scores assigned to each ship and then to rate each of these ships' personnel status based on your assessment of the component scores. Here are the 14 components:

1. (REENLIST) Reenlistment rate
2. (MAST) Nonjudicial punishment rate
3. (UAR) Unauthorized absence rate
4. (REENLIST-FT) First tour reenlistment rate
5. (TURNOVER) Long-term stability of personnel
6. (MANNING-NEC) Manning level ratio describing the proportion of billets manned by qualified persons according to the NEC data system
7. (PASS RATE) Percentage of persons taking rating exams who pass
8. (COURSES) Rate of correspondence course participation
9. (SACK TIME) Percent of time spent sleeping while off duty
10. (LEADERSHIP) Perceived leadership effectiveness by crew members
11. (DISCHARGE) Discharges other than honorable (including reference to special BuPers code)
12. (MAINTENANCE) Percentages of maintenance actions deferred due to insufficient manning or expertise (from 3M data system)
13. (SAT-PA) Satisfaction with present assignment (as measured by confidential questionnaire to sample of officers and enlisted)
14. (ENLIST-FIT) GCT + ARI + Education Level + Mean Semi-Annual Evaluation (all scores standardized) for all or for a sample of enlisted persons averaged across the ship.

Now that you have received a very basic outline of what we will be doing today, let's digress a bit and supply some background concerning this project. The Office of Naval Research (ONR) has asked Personnel Decisions, Inc. (our firm) to move toward developing a Naval Personnel Status Index. The

personnel status index (NPSI) is to be an indicator of the state of the human organization or of the personnel status of organizational units such as ships. As part of our NPSI development effort, we visited the Navy Postgraduate School in January to ask experienced officers to help us generate a list of components with good potential for contributing to NPSI measurement.

The result of the January workshops here was a list of 29 component measures thought to have potential for indicating the personnel status of ships. On the second day of the workshops, officers also provided evaluations of the "goodness" of each component along five criteria--Importance, Reliability, Generalizability, Accessibility, and Fudgeability. Analysis of these evaluations enabled us to reduce the number of components for further consideration to 14. These 14 components seem to represent a wide spectrum of possible measures for a comprehensive NPSI.

Broadly speaking, the purpose of the rating task today is to determine how officers use this information from components to assess the personnel status of ships. A more elaborate explanation of our purposes in this rating task would possibly affect your evaluations; consequently, we will defer those explanations until later. A "Debrief Package" describing more completely the purpose of this rating session will be made available to you after the session.

Let's now turn to the rating task and describe in detail what we need from you. First, you should have the following materials:

- a. a short questionnaire
- b. a page containing a definition of "personnel status" and a description of each of the 14 components
- c. 100 "ships" each described by 14 component scores
- d. five "practice ships" each described by 14 component scores
- e. a rating worksheet to record your evaluations

After you have checked your materials, please fill out the questionnaire and then get ready to begin the ratings. In order to insure that everyone interprets the rating task in a similar way, we will now discuss a few ground rules and some things to remember about the rating procedures.

- a. Component scores are based on a (1-9) scale where 1 = extremely unfavorable; 5 = average; and 9 = extremely favorable. It is very important that you interpret each component score along a favorable-unfavorable continuum. This will enable you to forget about the fact that "high" on some components is good, but on other components "low" is good. For example, a "high" reenlistment rate is "good," while a "low" nonjudicial punishment rate is "good." You need not worry about "high" and "low" in that sense. Instead, remember that the component scores relate only to how favorable ships stand along these components.

- b. When considering individual component scores for a ship, assume that each score was derived in a reasonably competent manner. Let's use two examples of components not included in the list of 14 in order to illustrate this point. For the component "Rate of Participation in Group Activities," you should assume that a reasonably good definition of "Group Activities" could be generated and that the measure of the rate of participation in these activities was done properly. For the component "Satisfaction with Shipmates," you should assume that the questionnaire designed to measure this component would be reasonably well constructed and that the persons aboard each ship being assessed would be properly sampled and would do a reasonably conscientious job of completing the questionnaire.

In other words, assume that the scores assigned to the ships within this rating task are reasonable estimates of the "true" scores for these components (to the limits of the various measures' capabilities to provide uncontaminated, "good" data).

- c. For your first pass through the ships, we suggest that you evaluate them at a rate of approximately one per minute. Some of you may be able to assimilate more quickly the information provided by the component scores and some may need longer to do a conscientious job with the ratings. However, our pretest suggested that one ship per minute is a reasonable rate to shoot for. (Of course, the first five to ten ships may go a little slower because you will still be comparatively unfamiliar with the components.)

If you understand points a, b, and c above, you are ready to begin evaluating the first of five "practice ships" in your packet. To get everyone off on the right foot, we will now describe the procedure for making personnel status evaluations. First, read carefully the definition of personnel status. It is essential that you have a clear idea of what is meant by this term. Then read the component descriptions to get their general definitions firmly in mind. Now remove Practice Ship 1 from your packet. Study the component scores for that ship. When you have formed a clear impression of the status of the personnel subsystem on that ship, record your rating (from 1 = extremely poor to 7 = extremely good) next to the appropriate blank on the rating work sheet. Do the same for practice ships 2-5. If you have questions about the task at that time, feel free to discuss them with us. If you have no questions, you are ready to proceed with the "for real" ships.

Here is the procedure to follow for these 100 ships. Instead of simply rating ships, we ask you to sort each ship into a category (from 1 = extremely poor, to 7 = extremely good) according to a specific plan. On the basis of your judgments about the personnel status of each ship, you should sort ships into categories to satisfy the following requirements:

<u>Personnel Status Category</u>	<u>Number of Ships to be Sorted Into the Category</u>
extremely poor - 1	2
poor - 2	8
somewhat below average - 3	24
average - 4	32
somewhat above average - 5	24
good - 6	8
extremely good - 7	2

To accomplish this sorting task, we ask you to comply with the following procedure. For your first pass through the ships, take about one minute to study the component scores for each ship. When you have decided on an overall personnel status rating for a ship, record your evaluation (1-7) on the work sheet according to the number code on the ship (e.g., N0 or TV); also, sort the "ship" (card containing the component scores) into a pile containing other ships assigned the same rating. That is, as you proceed through the ships, form seven stacks, one for each of the seven points on the rating scale. Please try to remain roughly within the guidelines describing the desired rating distribution as you make your evaluations. For example, you should not be rating every ship above average.

After you have sorted each ship into one of the seven scale categories and recorded all ratings on the work sheet, count the number of ships you have placed in each category. If your distribution of ratings is more than a few off from the desired distribution, resort ships to other categories until your distribution is exactly like or very similar to the distribution outlined earlier. This operation will require you to go back through certain piles to find ships that should be resorted. After you have made all changes, record them on the work sheet. This step completes the task.

Below we provide an outline or a "check list" of the procedures presented above:

- 1) Insure that you have all required materials
- 2) complete the questionnaire
- 3) read ground rules on pages 2 and 3
- 4) read carefully the definition of personnel status, the descriptions of the components, and the directions for completing the ratings and the sorting
- 5) rate the five practice ships
- 6) rate and sort the 100 ships
- 7) resort ships (if necessary) to comply with the desired distribution
- 8) record all rating changes on the work sheet

The next page contains four "ships" extracted from the rating task. Within the task, component order will be randomly presented for each ship; i.e., the order of components will generally change for each succeeding ship presented to a rater. Also, the 100 ships will be presented to each rater in random order.

SHIP NUMBER 001 (NO)
COMPONENT SCORES

REENLIST-FT	****4
UAR	***3
REENLIST	****4
ENLIST-FIT	****4
TURNOVER	*****6
LEADERSHIP	***3
SAT-PA	***3
MAST	**2
MANNING-NEC	****4
COURSES	***3
DISCHARGE	*****5
MAINTENANCE	****4
PASS RATE	****4
SACK TIME	*1

SHIP NUMBER 002 (NP)
COMPONENT SCORES

MAINTENANCE	****4
REENLIST	***3
ENLIST-FIT	**2
SACK TIME	*****5
MANNING-NEC	***3
DISCHARGE	*1
MAST	*1
TURNOVER	*1
PASS RATE	*****6
UAR	*1
COURSES	*****7
LEADERSHIP	****4
REENLIST-FT	***3
SAT-PA	*****5

SHIP NUMBER 003 (NQ)
COMPONENT SCORES

REENLIST-FT	*****9
MAINTENANCE	*****7
REENLIST	*****9
COURSES	*****9
PASS RATE	*****7
MANNING-NEC	*****5
SAT-PA	*****6
MAST	***3
LEADERSHIP	*****8
SACK TIME	*****9
TURNOVER	*****8
DISCHARGE	*****6
UAR	*****5
ENLIST-FIT	*****8

SHIP NUMBER 004 (NR)
COMPONENT SCORES

MANNING-NEC	*****6
REENLIST-FT	*****5
UAR	*****5
DISCHARGE	*****5
COURSES	*****8
SACK TIME	*****6
MAST	*****5
ENLIST-FIT	*****8
LEADERSHIP	*****6
REENLIST	*****7
MAINTENANCE	*****8
PASS RATE	*****9
TURNOVER	*****5
SAT-PA	*****6

Rating Worksheet

1 = extremely poor

4 = average

5 = somewhat above average

2 = poor

6 = good

3 = somewhat below average

7 = extremely good

	<u>Practice Ships</u>	<u>Personnel Status Rating (1 - 7)</u>		
	1	_____		
	2	_____		
	3	_____		
	4	_____		
	5	_____		
<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>		<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
NO	_____		ON	_____
NP	_____		OO	_____
NQ	_____		OP	_____
NR	_____		OQ	_____
NS	_____		OR	_____
NT	_____		OS	_____
NU	_____		OT	_____
NV	_____		OU	_____
NW	_____		OV	_____
NX	_____		OW	_____
			OX	_____

1 = extremely poor

4 = average

5 = somewhat above average

2 = poor

6 = good

3 = somewhat below average

7 = extremely good

<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
PN	_____
PO	_____
PP	_____
PQ	_____
PR	_____
PS	_____
PT	_____
PU	_____
PV	_____
PW	_____
PX	_____

<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
QN	_____
QO	_____
QP	_____
QQ	_____
QR	_____
QS	_____
QT	_____
QU	_____
QV	_____
QW	_____
QX	_____

<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
RN	_____
RO	_____
RP	_____
RQ	_____
RR	_____
RS	_____
RT	_____
RU	_____
RV	_____
RW	_____
RX	_____

<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
SN	_____
SO	_____
SP	_____
SQ	_____
SR	_____
SS	_____
ST	_____
SU	_____
SV	_____
SW	_____
SX	_____

Rating Worksheet (Continued)

3

1 = extremely poor

4 = average

5 = somewhat above average

2 = poor

6 = good

3 = somewhat below average

7 = extremely good

<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
TN	_____
TO	_____
TP	_____
TQ	_____
TR	_____
TS	_____
TT	_____
TU	_____
TV	_____
TW	_____
TX	_____

<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
VN	_____
VO	_____
VP	_____
VQ	_____
VR	_____
VS	_____
VT	_____
VU	_____
VV	_____
VW	_____
VX	_____

<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
UN	_____
UO	_____
UP	_____
UQ	_____
UR	_____
US	_____
UT	_____
UU	_____
UV	_____
UW	_____
UX	_____

<u>"Real" Ships</u>	<u>Personnel Status Rating (1 - 7)</u>
WN	_____
WO	_____

QUESTIONNAIRE

1. Your present rank _____
2. Your "specialty," if applicable; e.g., aviator, submariner, destroyer officer, etc. _____
3. List the types of ships you have served aboard

4. How familiar are you with the N&C data system?
____ very familiar; ____ familiar; ____ know only basically what it involves; ____ unfamiliar.
5. How familiar are you with the 3M data system?
____ very familiar; ____ familiar; ____ know only basically what it involves; ____ unfamiliar.
6. How would you assess your overall familiarity with administrative systems and indices kept in BuPers, aboard ship, and at various head-quarter levels?
____ very familiar; I have considerable experience with a wide variety of such systems and indices.
____ quite familiar; I have at least average familiarity with such systems and indices compared to an experienced LCDR or CDR.
____ familiar with most such systems and indices with the exception of the following areas:

____ not very familiar; I have relatively little experience with such systems and indices.

6/ DEFINITION OF PERSONNEL STATUS

Personnel Status is the component of operational readiness related to the personnel subsystem of a ship. It is that part of operational readiness or of potential unit effectiveness which is affected directly by the status of the personnel subsystem. Another way of thinking about a ship's personnel status is that the ship's overall readiness is a function of equipment capabilities and status, situational constraints, and the state of the human organization--we are focusing in on the last of these three broad factors related to operational readiness.

The Components

1. (REENLIST) Reenlistment rate
2. (MAST) Nonjudicial punishment rate
3. (UAR) Unauthorized absence rate
4. (REENLIST-FT) First tour reenlistment rate
5. (TURNOVER) Long-term stability of personnel
6. (MANNING-NEC) Manning level ratio describing the proportion of billets manned by qualified persons according to the NEC data system.
7. (PASS RATE) Percentage of persons taking rating exams who pass
8. (COURSES) Rate of correspondence course participation
9. (SACK TIME) Percent of time spent sleeping while off duty
10. (LEADERSHIP) Perceived leadership effectiveness by crew members
11. (DISCHARGE) Discharges other than honorable (including reference to special BuPers Code)
12. (MAINTENANCE) Percentages of maintenance actions deferred due to insufficient manning or expertise (from 3M data system)
13. (SAT-PA) Satisfaction with present assignment (as measured by confidential questionnaire to sample of officers and enlisted)
14. (ENLIST-FIT) GCT + ARI + Education Level + Mean Semi-Annual Evaluation (all scores standardized) for all or for a sample of enlisted persons averaged across the ship

Debrief Package

Responses you made on the rating task just completed will provide us with a model of your "policy" related to making judgments about personnel status. From your ratings, we will generate policies for individual raters and composite policies for groups of raters. The policy capturing analyses we are describing will yield an importance weight for each of the 14 components reflecting the effect scores on that component had on your overall personnel status evaluations. We will perform the same analyses on the group's composite ratings, allowing us to obtain importance weights appropriate for the whole group as well.

A central purpose for generating these importance weights is to provide a means for discovering which component measures experienced Naval officers weight heavily when assessing the personnel status of ships. In a sense, we will be formalizing and systematizing Naval officer wisdom by capturing the policy--i.e., obtaining component importance weights--of officers participating in this project. The way you and the other officers use and combine this information about ships' personnel status will yield important clues about what types of measures should be most legitimately included in an NPSI.

A related objective of this policy capturing rating task is to assess the degree of policy similarity across Naval officers. Data you provide will suggest either that officers use substantially different weighting systems in making judgments about ships' personnel status or that officers regard as important much the same kind of information when assessing the state of ships' personnel subsystems.

What will these results mean to the future of the project? As of now, we are not funded to proceed further with our study. However, if we receive money to go forward, we would first choose carefully those components which seemed most important on the basis of the policy capturing data. Then we would initiate a thorough study of the psychometric properties of one or more measures of each of these selected components. That is, we would focus on actually trying out various measures in an attempt to get the best indicator possible for each important component.

Once the operational measures were finally chosen (developed), the NPSI composite would be ready for a trial run. Basically, our research strategy would involve assessing the "construct validity" of the NPSI--i.e., finding out whether or not it was related to what it should be related to and insuring that it was not related to other measures or phenomena having little to do with the personnel status domain.

Once such an index was developed and validated, there would exist a number of potential applications. A couple of examples: An NPSI could be used to assess the effect of "people programs" instituted either Navywide or within lower level commands. Also, the index would be a valuable diagnostic tool for a person with command responsibilities. NPSI scores would suggest areas in the personnel subsystem which most needed his attention.

In closing, the Personnel Decisions' staff would like to thank you very much for participating in this research. We pledge to send one more newsletter describing the results of the policy capturing rating task you just completed. Hope to see you again--smooth sailing.