AD-750 761

ACTIVE SONAR TARGET DETECTION AND REPORTING: PERCEIVED CONSEQUENCES AND THEIR EFFECTS ON PERFORMANCE

C. Dennis Wylie, et al

Human Factors Research, Incorporated

Prepared for:

Office of Naval Research

June 1972

AD750761

TECHNICAL REPORT 1721-1
CONTRACT N00014-71-C-0257
NR 196-104

ACTIVE SONAR TARGET DETECTION AND REPORTING: PERCEIVED CONSEQUENCES AND THEIR EFFECTS ON PERFORMANCE

C. DENNIS WYLIE
ROBERT R. MACKIE

D D C
NOV 3 1972
C

PREPARED FOR
ENGINEERING PSYCHOLOGY PROGRAMS (CODE 455)
AND
NAVAL ANALYSIS PROGRAMS (CODE 462)
OFFICE OF NAVAL RESEARCH
DEPARTMENT OF THE NAVY
ARLINGTON, VIRGINIA 22217

JUNE 1972

HUMAN FACTORS RESEARCH, INCORPORATED
SANTA BARBARA RESEARCH PARK, GOLETA, CALIFORNIA

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

UNCLASSIFIED

Security Classification

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)
Human Factors Research, Incorporated
6780 Cortona Drive
Goleta, California 93017

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP
N/A

3. REPORT TITLE
ACTIVE SONAR TARGET DETECTION AND REPORTING: PERCEIVED CONSEQUENCES AND THEIR EFFECTS ON PERFORMANCE

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report

5. AUTHOR(S) (First name, middle initial, last name)
C. Dennis Wylie & Robert R. Mackie

6. REPORT DATE
June 1972

7a. TOTAL NO. OF PAGES 176
103 + Appendices

7b. NO. OF REFS
-21-

8a. CONTRACT OR GRANT NO.
N00014-71-C-0257

b. PROJECT NO.
NR 196-104

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
Technical Report 1721-1

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
None

10. DISTRIBUTION STATEMENT
Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES
None

12. SPONSORING MILITARY ACTIVITY
Engineering Psychology & Naval
Analysis Programs
Office of Naval Research

13. ABSTRACT

The objective of this research was to examine the effects of operational situations upon sonar operators' contact reporting behavior. Different operational situations have been observed to lead to different sonar contact reporting thresholds. This research was intended to develop a more complete and quantitative understanding of that effect. The aspects to be investigated were suggested by statistical decision theory. An attitude survey involving fleet sonar operators and destroyer officers was conducted to determine the subjective values of five decision-making variables with respect to seventeen scenarios depicting peacetime and wartime tactical situations. It was found that sonar contact reporting thresholds were principally determined by the perceived consequences of missed contacts and of delay in contact reporting. It was also found that the assessments of false contact consequences were very inconsistent, unlike the judgments with respect to the other decision variables.

The second phase of the research consisted of the conduct of a sonar detection experiment, in which the influence of the variable "command attention" on performance was measured to give an objective evaluation of the effects of psychological variables. It was found that a "high" level of command attention led to better sonar contact detection performance.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6801

I-A

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
False target False contact Command level decisions Detection Classification ASW						

DD FORM 1 NOV 65 1473 (BACK)
(PAGE 2)

I-B

ACTIVE SONAR TARGET DETECTION AND REPORTING:
PERCEIVED CONSEQUENCES
AND
THEIR EFFECTS ON PERFORMANCE

C. Dennis Wylie
Robert R. Mackie

Prepared for

Engineering Psychology Programs (Code 455)
and
Naval Analysis Programs (Code 462)
Office of Naval Research
Department of the Navy
Arlington, Virginia 22217

Prepared by

Human Factors Research, Incorporated
Santa Barbara Research Park
6780 Cortona Drive
Goleta, California 93017

Contract N00014-71-C-0257
NR 196-104

June 1972

II

ACKNOWLEDGMENTS

A study of this complexity requires the cooperation, suggestions, and assistance of many individuals and groups. We would particularly like to express our appreciation to the following:

Dr. Martin A. Tolcott, ONR, Washington;

Mr. J. R. Simpson, ONR, Washington;

Dr. Charles M. Harsh, ONR, Boston;

Captain Thomas H. Sherman, Commanding Officer, DESDEVGRU, and Staff;

Captain R. R. Law, Commanding Officer, Fleet ASW School, San Diego, and Staff;

and to the many DESLANT and DESPAC personnel who willingly gave not only their time but thoughtful consideration to the problem under study.

TABLE OF CONTENTS

Preceding page blank

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

SUMMARY

Objective

The objective of this research was to examine the effects of various operational situations upon sonar operators' contact reporting behavior. The act of reporting a contact constitutes the only operationally significant definition of "detection," since, in the absence of such a report, a ship has not made a "detection" and cannot react appropriately, regardless of the state of mind of a sonar operator. Different operational situations have been observed to lead to different sonar contact reporting behaviors. This research was intended to develop a more complete and quantitative understanding of that effect, to permit more accurate prediction of its influence on ASW performance, and to suggest means for modifying reporting behavior, if necessary, to optimize ASW performance

The Attitude Survey

The aspects to be investigated were suggested by statistical decision theory. This theory identifies variables which have an important impact on optimum decision making. For the sequential detection situation, these variables include: (1) the decision criterion, or threshold; (2) the consequences of missed detections; (3) the consequences of false detections; (4) the *a priori* probability for the presence of a target; and (5) the "cost" of additional observations. For the case of the human decision maker, it has often been postulated that the *subjective* values of these variables influence decision-making behavior in a manner similar to the mathematical model. The first phase of the research focused on obtaining the judgments of fleet sonar operators and destroyer officers regarding the subjective values of these decision-making variables for each of 17 operational scenarios, which depicted a wide range of peacetime and wartime destroyer operations. The individual

scenario descriptions were printed on cards and each of the personnel interviewed was requested to place the 17 cards in rank order according to sets of instructions directing attention to the five decision-making variables. With few exceptions, the CO, XO, operations officer, ASW officer, weapons officer, and the sonar operators of 20 destroyers were interviewed, involving a total of 99 officers and 119 sonar operators.

By appropriate psychometric scaling methods, the judgmental data obtained in the survey provided not only information concerning the rank order of the scenarios on the various decision-variable dimensions, but their separations along these dimensions as well, based on the variability in responses. The principal findings of the scaling of the tactical scenarios on these psychological dimensions were:

1. Sonar contact *reporting thresholds* were principally determined by the subjective assessments of the *consequences of missed contacts* and of the *consequences of delay in contact reporting* (i.e., the "cost" of additional observations). The judgments of these consequences for each scenario were essentially identical, and result from what we might descriptively call the "threat" inherent in the tactical situations. To a lesser degree, the reporting thresholds were also found to depend upon the judged *probability of making submarine contact*, and the judged *consequences of false contacts*. These relationships were expected.

2. The following result was unexpected. While assessments of the scenarios with respect to *reporting thresholds*, *missed contact consequences*, *probability of submarine contact*, and the *"cost" of observations* were very consistent across all personnel interviewed, the assessment of the *consequences of false contacts* was very inconsistent. It was found that 55% of the personnel ranked the wartime scenarios as having the *most severe* false contact consequences, 28% of the personnel ranked the wartime scenarios as having the *least severe* false contact consequences, and 17% of the personnel followed no pattern at all in ranking the scenarios with respect to false contact consequences.

This inconsistency of response is believed to result from the peculiar status of the "false contact" in ASW, owing to the active sonar target classification problem. False contacts are the most frequent occurrence in surface-ship ASW operations of most kinds, but no tactic or equipment exists which can easily cope with the resulting attack decision problem. Owing to the lack of any easy answer to the false contact problem, the problem itself has often been dealt with superficially, both in the training of tactical officers, and in various research efforts concerned with surface ship ASW.

It was found that the diverse and inconsistent attitudes toward *false contact consequences* influenced the respondents' reporting threshold evaluations, and it seems likely that these attitudes would affect decision-making behavior in actual ASW operations, as well. Therefore, we regard this finding as an indication of a significant problem area.

In addition to these results, the application of multidimensional scaling techniques to the survey data revealed what appears to be a "personal consequence" dimension to the assessments of the decision variables, in addition to the expected "tactical consequence" dimension; and the responses to a questionnaire related to problems in ASW lent further insight into these problem areas.

The Detection Experiment

The second phase of the research consisted of the conduct of a sonar detection experiment, in which the influence of the "operational situation" on performance could be measured to give an objective evaluation of the effects of psychological variables. The physical "threat" inherent in wartime situations cannot be reproduced in an experiment, of course, but there is one operationally meaningful variable that relates directly to perceived importance of the tactical situation, if not to the perceived threat per se. That variable is "command attention." The intent was to measure operator detection performance against a given set of signals in a setting which would first minimize apparent command attention; and measure performance of the same operators against the same stimuli in a setting in which maximum possible command attention would be brought to bear, thus creating two situations genuinely perceived by the operators to be of very different importance with respect to their detection performance.

The SQS-26 A-scan display was used to generate stimulus materials. This display is a memory-type device which presents a history of sonar pings in a static presentation. This static characteristic permitted the use of still photography to create a realistic representation of the sonar display, by the rear-projection of ping-by-ping sequences of color transparencies onto translucent display screens of the same size as the sonar display CRT. Seventy-two 6-ping sequences of color slides were prepared, forty-eight of which contained FM or CW targets with random strengths, locations, and motion, generated by the sonar's target simulator; and twenty-four of which contained only reverberation and noise produced by an apparatus constructed for this purpose. Presentation of these sequences in random order at a real-time rate resulted in a detection task of approximately three hours duration.

During the course of a week, 18 experienced SQS-26 sonar operators performed this 3-hour detection task. During this period the operators were asked to report any contacts which they felt they would report aboard their ship during *routine* operations. Responses were required only when they wished to report a contact, in which case a push button was to be actuated, and the target range, bearing, and other information were to be recorded in a contact reporting booklet. The setting for this "low command attention" phase was that of an experiment conducted for the Navy by civilian researchers, but without explicit consequences for either errors or error-free performance of the task. The origin of the stimulus materials was never discussed with the sonarmen and knowledge of results was never given.

In contrast, when the same sonarmen arrived during the subsequent week, a U. S. Navy Commander, in uniform, was present in the room. He made a statement to the men following this guideline:

> The exercise last week was, in reality, just a warm-up for the very important task I want you to perform today. In the interest of developing optimum tactics against the Soviet submarine threat, you are being asked to view a series of recently acquired data where it is known that a number of targets may be detectable. Security prevents my discussing the nature of these data in detail, but it is extremely important that every valid contact be identified as such. Our estimate of the threat clearly depends on the most accurate information we can get in this respect. Therefore, please report any contact that you feel qualifies as a "possible sub." Report at the earliest point in the sequence where you feel such a report should be made.

Following this statement, the men were given response booklets which had security classification markings (unlike the first week's response booklets), and they were requested to perform another 3-hour detection task (which in fact employed the same stimulus materials that were used in the first week).

Thus, the setting during the second week was quite different from that during the first. Every effort was made to make performance of the detection task appear relevant and consequential to the "operational Navy," via the "high" level of command attention focused upon it. The results of the experiment may be summarized as follows.

1. The "high" level of command attention resulted in a significant improvement in detection performance, compared to the "low" level, by reducing FM false reports by 50% while reducing FM correct reports by

only 4%; and by reducing CW false reports by 31% while increasing CW correct reports by 20%. The differential effect by signal type is attributed to the different psychophysical tasks presented by the FM and CW portions of the display.

2. Each sonar contact report was required to have a contact "quality" judgment, using "good," "fair," and "marginal" contact quality categories. It was found that these quality judgments had diagnostic value; that is, a higher percentage of the contacts judged to be in the "good" category turned out to be true target contacts (as opposed to false target reports) than in the "fair" category, and so on. The effect of the "high" level of command attention was to improve target discrimination performance in every contact quality category, for both types of signals. For example, for FM targets in the "good" category, the proportion of correct reports was increased from 74% for the "low" level of command attention to 84% for the "high" level. Similar improvements were seen in the other categories.

3. Each sonarman was required to estimate the range of each contact he reported. During the "high" level of command attention, a 19% improvement in target range estimation performance was observed, compared to the performance at the "low" level of command attention.

4. On a Receiver Operating Characteristic graph, the "high" level of command attention moved the CW operating point away from the operating point for the "low" level almost perpendicularly to the "chance" diagonal, in a direction of improved effective signal detectability; and the displacement of the FM operating point had approximately equal components of perpendicular movement, corresponding to improved detectability, and parallel movement, corresponding to a change of reporting criterion in the direction of increased "caution."

5. Based upon certain assumptions, the signal detection theory statistics d' and β were calculated. It was seen that, on the average, signal detectability was greater for the FM signals than the CW signals; and that more "strict" reporting criteria (i.e., larger values of β) were employed for FM signals than for CW. It was also seen that the "high" level of command attention brought about a

> 35% improvement in FM signal detectability, and a 64% improvement in CW signal detectability, compared to performance at the "low" level. The reporting criterion concomitantly increased by 109% in the direction of "caution" for FM signals, and by 34% for CW signals.

Thus, target detection performance was significantly better during the high command attention phase than the low command attention phase of the experiment. However, it is very important to stress that the contrast between actual peacetime (non-ASW exercise) and wartime target detection and reporting behavior would stand in even greater contrast. While the sonarmen were asked to report any contacts "which they felt they would report aboard their ship during routine operations" for the low command attention phase, the setting was obviously one of a detection experiment, which places the emphasis on making target detections, and implies that there are targets to be detected. Furthermore, there was obviously no immediate consequence for any false detections that might be made. In contrast, the actual routine peacetime operational setting carries with it a very low probability of making a valid target contact; and there is an immediate consequence to reporting false target contacts, in that they are brought to the attention of command. The results of the attitude survey, anecdotal evidence collected from the survey respondents, and the results of ASW exercise reporting performance compared to non-exercise contact reporting indicate that the routine peacetime sonar contact reporting threshold is typically more "strict" than that employed during the low command attention phase of this detection experiment. Thus, the effect of the high level of command attention upon ASW performance at the onset of war could be expected to bring about the improved "effective signal detectability" which we have seen to occur in the detection experiment; and to bring about an overall increase in the number of target reports, as compared to routine peacetime operations, by lowering the reporting threshold to the comparatively "free" values observed in both phases of the detection experiment.

Conclusions

In the operational sonar detection situation, the noise/reverberation background is an insignificant source of "false" responses, compared with that unfortunately large class of objects which give rise to "false contacts." "False contacts" produce real echoes, whose characteristics differ significantly from the noise background, principally in having a degree of persistence impossible of the true noise/reverberation sources. Particularly during peacetime operations, these false contact sources greatly outnumber actual submarine contacts. The results of the attitude survey, the detection experiment, and specific anecdotal evidence collected from the survey respondents indicate that the peacetime "solution" to this considerable incidence of false targets consists of a "low detection efficiency, strict reporting threshold" mode of operation, brought about largely by the typically low level of command attention focused upon the sonarman's peacetime role. The consequence of this mode of operation is to screen out false (and valid) contacts at the lowest hierarchical level, relieving the rest of the team (CIC, UB, and Command) of the burden *and the experience* of the ASW decision-making process.

In the event of war involving ASW, however, that burden is going to shift. The high level of command attention which will undoubtedly be focused upon the ASW role, and the presence of an actual threat situation, will lead to a "high detection efficiency, low reporting threshold" mode of operation. And this is a desirable and probably optimum mode of operation, for it opens the channel from the primary ASW sensor to the entire ASW team, permitting classification and attack decision making to function with all the resources that can be brought to bear by the individual surface ship, including inputs other than sonar. The difficulty, as we see it, lies with maintaining this optimum state for destroyer ASW operations. At the onset of war, the naturally large number of objects which can cause false contacts, the greater wartime inclination for the sonarman to detect and report them, and the inexperience of the rest of the team in dealing with them (and their heterogeneous attitudes toward them) will constitute a serious ASW decision problem. The seriousness of the matter is heightened by the typical peacetime conduct of destroyer operations, which often denies the ASW team the experience of reporting and prosecuting the false contacts which *do* exist in the ocean medium, in peace or war.

We fear that unless destroyer officer personnel appreciate the false contact problem in its fullest extent, apparent command dissatisfaction with classification performance during the transition to a genuine ASW role will, intentionally or unintentionally, influence the sonar operator in the direction of undesirably conservative detection and reporting behavior. And it is a discouraging fact that very many destroyer commanding officers have been provided with very little ASW training or practical experience.

We see the destroyer commanding officer as the key determiner of ASW team decision-making behavior during the transition to a wartime ASW role. Research under this contract is now focusing on the destroyer officer's probable responses to wartime ASW decision situations.

CHAPTER I

INTRODUCTION

The Problem

It is necessary for the Navy to make quantitative, detailed, and accurate estimates of expected performance in a wide range of threat situations, and particularly in accomplishing its potential wartime missions. The performance of active sonar systems is critical to many of those missions, so it is essential that objective procedures for predicting the performance of sonar systems with man in the loop be developed. It is also necessary to develop sufficient understanding to permit modification and optimization of that performance where it is found to be inadequate or non-optimum.

The performance of an active sonar system is greatly dependent upon the behavior of the sonar operator as well as upon the functioning of electronic circuitry. The latter is comparatively well-known and deterministic. However, even the most sophisticated contemporary sonar equipment essentially only prepares information for a far more complex processing step: the operator's detection decision. If this decision were not so complex, we would require the equipment itself to make it. Because this processing step is *not* described in any electronic blueprint, it has not traditionally been a candidate for quantitative analysis and predictive synthesis. Determination of detection performance has previously taken into account the physical characteristics of the signal, noise, and reverberation, but little attention has been given to the influence of *situational variables upon the operator's detection behavior* and the consequent effects on sonar performance predictions.

A precise definition of "detection" is, of course, essential. Because of the complexity of the active sonar target detection/reporting task, more than one definition might be possible. One frequently hears a distinction made between "detection" and "reporting"--the report of sonar to the bridge that a target has been "detected." It is also frequently heard that an operator's "reporting criterion" is higher than his "detection criterion," so that he does not report all that he "detects." While these observations indicate that "detection" might be defined separately from "reporting," they also reflect awareness of, and concern for, an important fact: that sonar operators "filter" information in transmitting it from the sensor to those in command. The exact nature of such "filtering" has not been quantitatively known.

Such uncertainty is cause for concern, since *the act of reporting a contact constitutes the only operationally significant definition of "detection."* In the absence of such a report, *a ship* has not made a "detection," and cannot react appropriately, regardless of the state of mind of a sonar operator. In order to best estimate expected performance of the Navy in facing any ASW situation, operator *reporting* behavior must be known for that situation. Often, detection performance has been evaluated in experimental situations in which no attempt was made to relate the operator's "experimental detection criterion" to the criterion he would use for sonar contact *reporting* during operational situations. The "recognition differential" determined by such an experiment may be useful in making equipment performance comparisons, but it bears an uncertain relationship to the expected performance of a ship in facing an ASW threat. Knowledge of reporting behavior necessarily encompasses intermediary detection behavior, but the converse is certainly not true. It has been the basic objective of the research described in this report to be concerned with sonarmen's target reporting behavior and the influence of the operational milieu upon it.

The Approach

Our approach to understanding the sonar reporting decision has been guided by statistical decision theory. Statistical decision theory provides normative models for decision making. When various decision parameters and explicit decision goals are specified, the theory prescribes mathematically optimum procedures for decision making. The subset of the mathematical theory dealing with binary decisions (typically, the decision concerning the presence or absence of a signal in noise) has been particularly well developed in the context of electrical communication. Subsequently, the "theory of signal detectability" has come into wide use as a mathematical model for the human as a detector in psychophysical tasks. While the appropriateness of the assumptions necessary to make this application are not without controversy, the approach has rendered the singular, uncontroversial contribution of drawing well-deserved attention to previously neglected aspects of human detection performance: the "false alarm" error; the influence of signal probabilities; and the influence of subjective costs (or values) of various decision outcomes. There is now a large literature reflecting great attention to these aspects of decision making for psychophysical detection tasks. Green and Swets (1966) give an excellent introduction to the application of signal detection theory to psychophysics; Jeffress (1969) discusses these matters in the context of sonar detection.

However, most laboratory psychophysical tasks, while bearing somewhat questionable relationships to the underlying assumptions of the common signal detection theory model, because of their complexity, are *themselves* simplistic compared to most "real-world" detection tasks. Two things in particular distinguish the sonar detection task from common psychophysical tasks: its "vigilance" aspect; and its "sequential" nature. In the psychophysical task, the "basic human perceptual mechanism" is usually the principal object of study, and it is usually the desire to have experimental results unconfounded with the effects of low signal probability, varying degrees of arousal and motivation, prolonged work periods, and other sources of variation inherent in the typical operational task of searching for infrequent, difficult-to-detect signals over relatively long periods of time. The presence of these effects, organic to the sonar detection situation, strain the assumptions underlying the application of the most common signal detection theory model (i.e., that which assumes equal-variance Gaussian noise and signal-plus-noise distributions). For a discussion of the application of this model to vigilance tasks, see Broadbent (1971).

Further complicating the application of the simple model is the sequential nature of the sonar detection task. The sonar operator is not presented with a clearly defined "opportunity" in which he must respond either positively or negatively regarding the presence of a signal; rather, the situation is one in which he may either decide to report a contact, or to gather more information. Indeed, this particular situation even omits one of the alternatives of Wald's sequential decision model (Wald, 1947), in that no "rejection" decision is made overtly; the operator either reports, or continues observing. Birdsall and Roberts (1965) recently have theoretically extended Wald's sequential analysis to include costs and probabilities, producing what may be the most appropriate fabric for a model of sonar contact reporting; however, this application has not yet been attempted.

If a precise yet satisfactory theoretical description of a task as complex as sonar detection is wanting, however, a more general but valuable insight from statistical decision theory is not, for the theory directs our attention toward the potentially important variables of the decision-making process. These variables are listed below in the context of the sonar reporting decision.

1. *Contact reporting criterion or threshold:* it is postulated that a sonar contact is reported only when confidence in that contact, formulated from observation of the sonar displays, exceeds a "reporting threshold," which itself is a function of the other decision variables, listed below.

2. *Missed contact consequences:* the perceived consequences, or "cost," of an error of omission.

3. *False contact consequences:* the perceived consequences, or "cost," of an error of commission.

4. *Likelihood of making contact:* the subjective *a priori* probability of a valid sonar contact.

5. *"Cost" of additional observations:* the perceived consequences of delaying sonar contact reporting, to obtain additional information.

Two other variables of the general decision situation were felt inappropriate to the sonar contact reporting task. These are the "values" of the correct outcomes of the detection decision: "correct detection," and "valid rejection." It was found through pilot survey work that the "value" of a correct detection was a difficult concept for the sonar operators to evaluate; and the valid rejection is not appropriate because it is accompanied by no overt response from the sonar operator (assuming, of course, that the contact has not already been reported).

Theory (and common sense) suggests that variables 1 through 5 are important decision-making variables. As applied to human decision making, these variables may be regarded as constituting *psychological dimensions*, and subjective values on these dimensions for a given decision situation may influence decision making. The first objective of this research was to conduct an attitude survey and apply psychometric scaling methods to determine how *realistic tactical scenarios* are evaluated on these dimensions *by the decision makers themselves*--a representative sample of destroyer sonar operators.

In addition, the evaluation of the scenarios by destroyer officers was desired. This was motivated by the fact that the theoretical detector can only be said to be "optimum" with respect to the *given* consequences and *a priori* probabilities. *The determination of those consequences and probabilities is outside the scope of optimum detection theory.* Therefore, even if the sonar operator behaved precisely as an ideal detector, that behavior would only be optimum for the decision variable values *as he perceives them.* Nothing guarantees that his evaluation of an operational situation, in terms of consequences and probabilities, is "optimum" in any sense. A definition of optimality for these variables would, indeed, be an elusive thing. However, it is in part the judgment of just such elusive things as the consequences of missed detections and false alarms, and of the probability of meeting one's adversary, which characterize the responsibilities of command. Thus, the survey was designed to include a representative sample of destroyer officers, as well as sonar operators, to permit comparisons between these groups. The survey method, procedure, and results are described in detail in the next chapter.

Scaling realistic tactical scenarios in a psychological decision-variable space was expected to result in a unique and valuable contribution to understanding the nature of the important theoretical decision variables for the specific, practical task of sonar contact detection and reporting. However, the quantitative influence that situational variables exert upon actual sonar contact reporting behavior cannot be determined solely from judgmental data; the relationships must be verified experimentally. Many laboratory psychophysical detection experiments have shown that *specific* instructions to subjects regarding their reporting criteria can influence those criteria significantly; and this has been shown for a simulated sonar detection task as well (Kostoff and Montgomery, 1970). Other laboratory experiments have investigated the influence of situational variables on signal detection and reporting behavior through the less direct technique of providing "payoff matrices" to the subjects, explicitly specifying decision outcome "costs" and "values," often involving small monetary rewards (e.g., Williges, 1971). A sonar detection experiment has also been conducted employing the "payoff matrix" technique (Rizy, 1972). These studies typically show that "payoff matrices" have some influence on *reporting criteria* ("β" in the usual signal detection theory model) and none on *detection efficiency*, or signal detectability for the man-machine combination ("d'" in the usual model).

We feel that "payoff matrices," involving either monetary rewards or simply abstract incentives, provide very poor approximations to the motivating forces central in the question of "peacetime" versus "wartime" sonar contact reporting behavior. We have little doubt that "reporting thresholds" can be directly influenced by specific instructions, but we suspect that the apparent invariance of "detection efficiency" in these experiments was due to the relative impotence of the "payoff matrices" employed as experimental variables. While it is not (humanely) possible to reproduce the physical threat accompanying decision making in war, we were convinced that a sonar detection experiment *could* be conducted in a context genuinely perceived by Navy sonar operators to be of great importance, and of particular relevance to their principal occupational task.

Therefore, the second objective of the research reported here was to conduct a sonar detection experiment involving a realistically potent variable, which we have called "command attention." The method, procedure, and results of this experiment are described in detail in the chapter entitled "The Detection Experiment."

The conclusions we have drawn from this research are presented in the final chapter.

CHAPTER II

THE ATTITUDE SURVEY

Method

The objective of the survey was to obtain the judgments of Naval officers and fleet sonar operators regarding the positions of several tactical scenarios along certain psychological dimensions selected for their relevance to the sonar contact reporting decision. Five psychological dimensions were selected, and 17 tactical scenarios were selected to be scaled on these dimensions. Because of the relatively large number of dimensions and scenarios, the method of rank ordering was selected to permit obtaining the judgmental data within reasonable time constraints.

The scenarios, ranking instructions, and procedure are described in the next sections. The psychometric scaling methods used to evaluate the resulting judgmental data are described in the "Scaling Technique" section, and the derived scale values in the "Results" section. In addition to the ranking tasks which the respondents were requested to perform, each was asked to complete a questionnaire which contained questions concerning the respondent's background and viewpoints related to sonar contact reporting. The biographical data thus obtained are discussed in the "Respondents" section, and the responses to the sonar contact reporting questions are discussed in the "Results" section.

Scenarios

Seventeen tactical scenarios were written for use as stimuli for the survey. Since the objective of the survey was to obtain judgments which would have the closest possible connection to actual or anticipated operational situations, our principal guideline in the composition of these scenarios was to depict realistic situations representing a broad spectrum of decision variables. No attempt was made to construct scenarios to achieve predetermined levels of the various decision variables in all possible combinations. Stimuli generated in this way may be of academic value in investigating decision making, but treating the various aspects of tactical situations as *independent variables* very frequently leads to *unrealistic scenarios*, and the resulting judgmental data stand in a questionable relationship to existing or anticipated ASW missions. Our approach was to treat decision-making variables as dependent functions of the specific tactical situation, and our goal was to determine the subjective values of these variables, as judged by the potential decision makers themselves, for operational situations of the greatest practical interest.

Each of the 17 scenarios is shown in Appendix A exactly as it was presented to the survey respondents on individual 3" x 5" cards. The scenarios are numbered, and listed in numerical order, but the numbers were assigned to the scenarios randomly.

Ranking Instructions

Each respondent in the survey sample was asked to rank order the 17 scenarios five times, once according to each of the five sets of ranking instructions. Each set of instructions was designed to direct their attention toward a particular aspect of the sonar detection/reporting decision. The rationale for each of these sets of instructions, and the instructions themselves, will be given in turn. There exist slight differences between the instructions given to officers and to sonar operators. The instructions shown below are the operators' versions; the complete sets of instructions for both officers and operators are given in Appendix B.

Sonar Contact Reporting: The first set of instructions given to each respondent directed his attention toward the confidence level, or degree of certainty, felt necessary in order to make a sonar contact report. Presumably, the reporting decision involves weighing the physical evidence of a contact, as displayed by the sonar, against some confidence-level criterion for reporting. In the usual decision-theoretic model, the result of observation is expressed as some monotonic function of the likelihood ratio, which is then compared to a criterion derived from decision goal considerations, generally involving probabilities and decision consequences, to arrive at a decision. The intent of the first set of ranking instructions was to cause the respondents to evaluate and compare the decision criteria (analogous to those of the decision-theoretic model) which they felt they would employ in the various scenarios.

SONAR CONTACT REPORTING

> For the purposes of this part, please imagine that you are standing a sonar watch aboard your ship, that you are directly operating and observing your sonar, and that *you and you alone* will make the decision to report sonar detections to the bridge. Your certainty of a contact depends on many things, such as echo quality, consistency, strength, and so on, and you can be more sure of some contacts than others. If you were to see/hear a very strong echo which showed obvious submarine target cues or characteristics, you could report "possible sub" to the bridge with little doubt or uncertainty concerning the contact. On the other hand, if you were to see/hear an "echo" which was very weak, inconsistent, and lacking in cues, you might not be sure that you actually have a contact. Your decision to report such a questionable "contact" to the bridge might depend upon the operational situation--for example, you might be more likely to report such contacts during wartime ASW operations than you would in non-ASW peacetime situations. Please carefully read and consider each of the situations described on the cards, and place the cards in rank order in front of you so that the situation in which *YOU FEEL YOU WOULDN'T NEED TO BE VERY SURE OF A CONTACT TO REPORT "POSSIBLE SUB" TO THE BRIDGE* is at the top, downward situation-by-situation, to

> the situation for which *YOU WOULD WANT TO BE PRETTY SURE OF A CONTACT TO REPORT*. After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

Consequences of Missed Contacts: The second set of ranking instructions given each respondent directed his attention toward the consequences of errors of omission, or what are commonly referred to in statistics as "errors of the second kind"; that is, the consequences of missed contacts.

> CONSEQUENCES OF MISSED CONTACTS
>
> For the purposes of this part, please imagine that you are the sonar operator and that a "contact" briefly caught your attention, but was so weak, intermittent, and lacking in cues that you had little confidence that it actually *was* a contact and did not report it to the bridge as "possible submarine." At least part of the time, such contacts could actually be caused by submarines. If it *were* actually a submarine, it is a missed contact situation, which can have a variety of consequences. Your ship may lose points during an exercise, in time of war your ship or those in company may be torpedoed, the submarine may move out of range and never re-appear, with its existence remaining unknown, and so on. The exact consequences of a missed contact may be different for different people, and may depend upon the particular operational situation. *Please imagine that you have actually missed a contact.* Carefully consider what the consequences of a missed contact might be *as you see them* for each of the operational situations described on the cards, and place the cards in rank order in front of you so that the situation whose missed contact consequences are *MOST SEVERE, MOST UNDESIRABLE*, and/or *MOST UNPLEASANT* to you is at the top, downward situation-by-situation to that whose missed contact consequences are *LEAST SEVERE, LEAST UNDESIRABLE*, and/or *LEAST UNPLEASANT*. After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

Consequences of False Contacts: The third set of instructions directed the respondent's attention to the consequences of errors of commission, or errors of the "first kind"; that is, to the consequences of false contacts. The perceived consequences of errors of this kind were expected to be particularly interesting. Operators are taught to "report everything," but anecdotal evidence indicates that operators are reluctant to commit "false contact" errors, at least in some situations, to the extent of significantly affecting their target reporting behavior.

> CONSEQUENCES OF FALSE CONTACTS
>
> For the purposes of this part, please imagine that you are the sonar operator and that you have reported a contact to the bridge as "possible submarine." At least part of the time, a contact reported as "possible sub" turns out to be non-submarine. This situation constitutes a false contact, and can have a variety of consequences. Fuel and/or weapons may be expended, the ship may leave a position in a screen unguarded, the ship's captain may have to be awakened, and so on. The exact consequences of a false contact may be different for different people, and may depend upon the particular operational situation. *Please imagine that you have actually reported a false contact.* Carefully consider what the consequences of reporting a false contact might be *as you see them* for each of the operational situations described on the cards, and place the cards in rank order in front of you so that the situation whose *FALSE CONTACT CONSEQUENCES WOULD BE MOST SEVERE, MOST UNDESIRABLE,* and/or *MOST UNPLEASANT* is at the top, downward situation-by-situation to that whose *FALSE CONTACT CONSEQUENCES WOULD BE LEAST SEVERE, LEAST UNDESIRABLE,* and/or *LEAST UNPLEASANT.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

Likelihood of Making Contact: The fourth set of instructions directed the respondent's attention to the *a priori* probability of contact for each of the scenarios. Expectation, or subjective prior probability, is an important variable in decision-theoretic models of human decision making, and in actually observed decision behavior, and was expected to be an important variable in sonar contact reporting.

LIKELIHOOD OF MAKING CONTACT

> The likelihood of actually making contact with a submarine varies from situation-to-situation. Please carefully consider each of the situations described on the cards, and place the cards in rank order in front of you so that the situation you judge *MOST LIKELY TO RESULT IN SUBMARINE CONTACT DURING ONE SONAR WATCH* is at the top, downward situation-by-situation to that *LEAST LIKELY TO RESULT IN SUBMARINE CONTACT DURING ONE SONAR WATCH.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

Response Time: The final set of instructions called attention to the consequences of delay in sonar contact reporting. It was expected that scaling the scenarios according to the importance of response time would provide insight regarding the "cost of observation" variable which is important in sequential decision situations, including the sonar contact reporting task.

RESPONSE TIME

> A quick detection and report to the bridge may be more important in some situations than in others. Please carefully consider each of the situations described on the cards and place the cards in rank order in front of you so that the situation for which *DELAYING CONTACT REPORTING WOULD BE WORST* is at the top, downward situation-by-situation to that for which *DELAYING CONTACT REPORTING WOULD BE LEAST BAD.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

To identify any problems in our approach, a pilot survey was conducted at the Fleet ASW School, San Diego, involving ten sonar technicians and five Naval officers. The results of this pilot survey were very satisfactory, leading only to minor modifications in the wording of the instructions. The instructions we have just described are those of the final form, used in the main attitude survey.

Procedure

During June and July of 1971, project personnel visited the ports of Norfolk, Virginia; Newport, Rhode Island; Long Beach, California; and San Diego, California. A total of 25 destroyers and destroyer escorts were visited. Prior arrangements had been made to have two groups assembled for interview on each ship, with the compositions shown in Table 1.

TABLE 1

OFFICER AND ENLISTED GROUP COMPOSITION

Officer Group	Enlisted Group
a. Commanding Officer	a. Leading Sonar Technician
b. Executive Officer, Weapons Officer, or Operations Officer (two of the three)	b. Three Sonar Watch Supervisors
c. ASW Officer	

The officer and enlisted groups were assembled *separately* (usually, the officer group was interviewed in the wardroom and the enlisted group on the enlisted messdeck), but interviewed simultaneously by two of our personnel.

After a rather general introduction, each participant was given a shuffled deck of scenarios, the "introduction" page of the instructions, and the "Sonar Contact Reporting" page. The introduction (see Appendix B) explained the general nature and purpose of the survey, and contained a biographical questionnaire which was completed by each subject. The "Sonar Contact Reporting" ranking instructions have been described; each subject performed this ranking first. Sufficient space was provided for each person to lay out the 17 scenario cards, in order to facilitate the ranking process. When the ranking was completed, the scenario numbers were recorded by the subjects on the back of the ranking instruction sheets. After each respondent had completed the ranking according to the first set of instructions, the instructions were taken from him, and the next set of instructions and a reshuffled deck of scenarios was given him. Thus, no subject was able to make direct comparisons among either the sets of instructions, or their own responses. Each subject was given the ranking instructions in the following order: Sonar Contact Reporting; Consequences of Missed Contacts; Consequences of False Contacts; Likelihood of Making Contact; and Response Time. When each respondent had completed the ranking task, he was given a short questionnaire to complete concerning destroyer ASW. Approximately two hours were allowed to perform these tasks, which was sufficient in most cases.

The Respondents

A total of 99 Naval officers and 119 fleet sonar operators were included in the main survey sample. The composition of this sample is shown in Table 2. In general, considerable interest in the survey was expressed by the respondents, and the cooperation we received was very satisfactory. Several destroyer commanding officers requested, and were given, scenario decks for use in subsequent discussions of the various considerations raised by the survey.

TABLE 2

COMPOSITION OF SURVEY SAMPLE

OFFICERS		ENLISTED	
By Rate/Rank:		By Rate/Rank:	
CDR	18	STCS	2
LCDR	22	STCM	3
LT	29	STC	17
LTJG	19	ST 1	13
ENS	9	STG 2	54
CWO 2	2	STG 3	28
TOTAL	99	STG SN	2
		TOTAL	119
Officers by Billet:		Enlisted by Primary Duties:	
CO	19	Div. CPO	10
XO	18	Sonar Spvr.	14
Ops	18	Sonar Oper.	34
ASW	22	Sonar Maint.	46
Weapons	20	F/C Maint.	15
CIC	1	TOTAL	119
Engnr	1		
TOTAL	99		

The Scaling Technique

Mathematical models for decision making involve values, costs, and prior probabilities to deduce the expected value of decision outcomes, and to identify optimal decision-making processes. In human decision making, these considerations of costs and probabilities may also be taken into account, though very often numerical values for them are not available to the decision maker. Even when numerical values are available, however, their impact on decision making may depend upon the way the decision maker perceives them. It seems reasonable (and has been experimentally demonstrated for some decision situations) that "subjective" values, costs, and probabilities are important in human decision making. These subjective values are not necessarily the same as those measured on any objective scale; they represent values on a psychological scale as judged by the human decision maker. Statistical decision theory directs our attention toward the important variables of the rational decision-making process; However, the importance of these variables in human decision making is a matter for empirical verification.

In the case of signal detection, including the sonar detection process, decision theory suggests that six variables might be of particular influence: the "values" of the two correct decision outcomes, *correct report* and *correct dismissal*; the "costs" of the incorrect outcomes, *false alarm* and *false rejection*; the "cost" of *additional observations* (because the sonar situation is a sequential process); and the *a priori probability* for the presence of the signal. We postulate that every tactical situation, including the 17 represented by our scenarios, has some value on a psychological continuum with respect to each one of these considerations. The relationship of the ranking instructions to these considerations has been previously explained; the objective of the survey was to determine where the 17 scenarios fell on five of these psychological continua.

How can one determine the position of a scenario on one of these psychological dimensions? The most obvious way, perhaps, is the "quantitative judgment" method. A judge (a Naval officer or sonar operator) could be asked to estimate the *scale value* of each scenario along each psychological dimension; or to estimate the *distance* between pairs of scenarios on these dimensions; or to make *ordinal judgments concerning distances* between pairs of scenarios (e.g., the distance from "A" to "B" is greater than the distance from "C" to "D"). Each of these tasks supposes that *distances* between scenarios can be judged. This method requires considerable time and thought on the part of the judges, presuming that they can perform the task at all. Particularly with the present task, it was felt that the judges could not perform in a meaningful way with this method in any reasonable amount of time.

Another approach to identifying the scale values of the scenarios on the psychological continua, one which requires less of the judges in time and effort, and particularly in the assumed ability to make "distance" judgments, is the "variability judgment" method. In this method, each judge is asked to make only *ordinal* judgments concerning scenario scale values. For example, he could be asked, "which is greater (with respect to the pertinent property), scenario 'A' or scenario 'B'?"; or, he could be asked to place the scenarios in *rank order* (with respect to the pertinent property). Although each judge makes only ordinal judgments, *distances* between scenarios may be derived from the *variability* in the ordinal judgments among judges, based upon the following hypothesis: when two scenarios are not widely separated on a psychological continuum, there is likely to be less agreement regarding their order (that is, which is greater) than if the scenarios *are* widely separated. The assumption is that the relation between *agreement concerning the order of two scenarios* (measured, for example, by the proportion of judges forming the majority opinion) and *the distance between the scenarios* on the psychological continuum is monotonic. Therefore, for example, if 95% of the judges feel that "A" is greater than "B," but only 60% feel that "C" is greater than "D," the implication is that "A" and "B" are further apart on the psychological continuum (and, therefore, more easily distinguished) than "C" and "D."

This hypothesis is a very reasonable one and is widely accepted as a rather unrestrictive assumption. The next step necessary in the conventional method of scaling from variability in ordinal judgments requires a somewhat stronger assumption. How, exactly, does one transform variability in the ordinal judgments into distance on the psychological continuum. The first assumption asserts that this transformation is monotonic; but what is the exact relationship? A widely used mathematical model of this relationship was first presented by Thurstone (Thurstone, 1927).

Thurstone postulated that when a given stimulus (e.g., a scenario) is presented to a judge, it gives rise to what he called a "discriminal process," represented by a specific scale value on the psychological continuum of interest; that the value of the discriminal process resulting from a single presentation of the stimulus is a random variable with a normal, or Gaussian, distribution; and that the mean and the standard deviation of this normal distribution are to be taken, respectively, as the *scale value* of the stimulus, and its *discriminal dispersion*. In other words, at each presentation of the stimulus, its value is "judged" along the pertinent dimension, and each judgment reflects "error" from the "true" scale value, an "error" which is normally distributed about the "true" value. In this model, no attempt whatever is made to account for the sources of this dispersion, in physical, physiological, or psychological specifics. Such specificity is unnecessary for this model; on the contrary, because there are certainly *many* sources of judgmental error, appeal is made to the central limit theorem of mathematical statistics in justifying the Gaussian distribution assumption. In any event, the model is just that; it must be subjected to empirical verification.

Now, in the "variability judgment" method of scaling, we do not deal with direct estimates of scale values, but rather with variability in ordinal judgment. What are the implications of Thurstone's model for these? In making an ordinal judgment concerning two stimuli "A" and "B," we can say that the judge attends to the instantaneous *difference between the discriminal processes* resulting from stimulus "A" and stimulus "B." If this discriminal difference is positive (subtracting "A" from "B"), stimulus "B" is judged to be greater than the stimulus "A"; and vice versa if the difference is negative. Obviously, since the discriminal process resulting from a particular exposure to a stimulus is a random variable, the discriminal difference is also a random variable. And, because the distributions of the discriminal processes of both stimuli "A" and "B" are postulated to be normal, the distribution of the discriminal difference is itself normal. We can most conveniently express these relationships in mathematical notation, as follows.

Let the "discriminal process" arising from stimulus "i" be denoted the random variable "X_i." In the Thurstone judgment model, the probability distribution of "X_i" is assumed to be normal, and thus completely defined by its mean and standard deviation. Let the mean of this distribution be denoted by "S_i" and the standard deviation by "σ_i." In the model these parameters represent, respectively, the "true scale value" of stimulus "i," and the "discriminal dispersion" of stimulus "i." Thus, to use the convenient mathematical notation wherein the normal distribution with mean "μ" and variance "σ^2" is denoted "$N(\mu,\sigma^2)$," the distribution of the discriminal process X_i is $N(S_i,\sigma_i^2)$.

Now, let the "discriminal difference" between stimuli "i" and "j" be denoted by "Δ_{ij}." By definition, $\Delta_{ij} = X_j - X_i$. Therefore, since X is a normal random variable, Δ is also a normal random variable, which, by the convenient properties of the normal distribution (we assume X_i and X_j to be uncorrelated) has itself the normal distribution $N(S_j - S_i,\ \sigma_i^2 + \sigma_j^2)$.

Let the mean of this distribution, "$S_j - S_i$," be denoted by "D_{ij}," the "true scale difference" between the stimuli in the Thurstone judgment model. The standard deviation of the discriminal difference is $(\sigma_i^2+\sigma_j^2)^{1/2}$.

To return to the ordinal judgment concerning stimuli "A" and "B," we saw that if the discriminal difference $\Delta_{AB} = X_B - X_A$ was positive for a particular observation, or judgment, stimulus "B" was judged to be greater than stimulus "A." What is the probability that this event will occur? It is the probability that the event $\Delta_{AB} > 0$ will occur, which is easily calculated, since Δ is normal with known parameters. Since Δ_{AB} is $N(D_{AB}, \sigma_A^2+\sigma_B^2)$, we know from the properties of the normal distribution that the random variable $Z = (\Delta_{AB} - D_{AB})/(\sigma_A^2+\sigma_B^2)^{1/2}$ is $N(0, 1)$. Because tables of the distribution function of this random variable (the normal deviate, or standard score) are commonly available, it will be convenient to ask the question "what is the probability that $\Delta_{AB} > 0$?" in terms of Z, as follows.

If: $Z = (\Delta_{AB} - D_{AB})/(\sigma_A^2\ \sigma_B^2)^{1/2}$

Then: $\Delta_{AB} = Z(\sigma_A^2+\sigma_B^2)^{1/2} + D_{AB}$

Therefore if: $\Delta_{AB} > 0$

then: $Z\ (\sigma_A^2+\sigma_B^2)^{1/2} + D_{AB} > 0$

or: $$Z > \frac{-D_{AB}}{(\sigma_A^2+\sigma_B^2)^{1/2}}$$

Thus, the probability that $\Delta_{AB} > 0$ is equal to the probability that $Z > -D_{AB}/(\sigma_A^2+\sigma_B^2)^{1/2}$, which is easily found in the tables. If $\Phi[z] = \Pr[Z \leq z]$ (i.e., the normal distribution function)

$$
\begin{aligned}
\text{then } \Pr[\Delta_{AB}>0] &= \Pr[Z>-D_{AB}/(\sigma_A^2+\sigma_B^2)^{1/2}] \\
&= \Pr[Z\leq D_{AB}/(\sigma_A^2+\sigma_B^2)^{1/2}] \\
&= \Phi[D_{AB}/(\sigma_A^2+\sigma_B^2)^{1/2}] \\
&= \Phi[z_{AB}]
\end{aligned}
$$

where $z_{AB} = D_{AB}/(\sigma_A^2+\sigma_B^2)^{1/2}$. Thus we see that the probability that stimulus "B" is judged greater than stimulus "A" is just the area under the normal curve up to the point z_{AB}, which is a standard score that represents the "true scale difference" between stimuli "A" and "B" in units of the standard deviation of the discriminal difference.

Now, we wish to experimentally derive (from ordinal judgments) information concerning *distances* between scenarios. This may be done in the following way. When stimulus "i" is compared to stimulus "j" by each of N judges, the experimental observable is the proportion "p_{ij}" of the N judgments asserting that stimulus "j" is greater than stimulus "i" on the psychological scale of interest. According to the Thurstone model, the probability of deciding "j>i" for one judgment is $\Phi[z_{ij}]$, so the expected number of such decisions in N judgments is $N\Phi[z_{ij}]$, and, therefore, the *expected proportion of such decisions in N judgments is* $N\Phi[z_{ij}]/N = \Phi[z_{ij}]$. We regard the experimentally observed proportion of such decisions as an estimator of the mathematically expected proportion:

$$p_{ij} = \Phi[z_{ij}]$$

Therefore, we may estimate z_{ij}:

$$z_{ij} = \Phi^{-1}[p_{ij}]$$

where $\Phi^{-1}(p)$ gives the ordinate (i.e., the standard score) corresponding to the area "p" beneath the normal curve.

The estimated z_{ij} for each pair of stimuli represents an estimate of the "true scale difference," D_{ij}, between those stimuli *scaled by the factor* $(\sigma_i^2 + \sigma_j^2)^{\frac{1}{2}}$, *the standard deviation of the discriminal difference.* If $\sigma_i = \sigma_j = \sigma$ for all "i" and "j," the scale factor becomes $\sqrt{2}\sigma$. One of the common simplifying assumptions in employing Thurstone's scaling model is this assumption (that the discriminal dispersions for all stimuli are equal), making the standard deviation of the discriminal difference identical for all pairs. If this is so, we may construct an *equal interval* scale by placing the scenarios so they are separated by the experimentally determined z_{ij}, and be assured, for example, that if scenario "A" is twice as far from scenario "C" as is scenario "B" on the scale so constructed, this distance relationship also holds on the psychological scale of interest. This property, of course, defines an equal interval scale, and is independent of the selection of an origin and unit for the scale. *The origin and unit of the scale are not defined by paired comparison or rank-ordering judgments*, which underlie the present technique, and, therefore, must be derived from other considerations (if at all). This is actually an advantage of the technique for scaling the tactical scenarios, because the selection of an origin and objective unit for "missed contact consequences," for example, can be very difficult (and very variable). Thus, any scaling method based upon the supposed ability of individual Naval officers or sonar operators to evaluate "consequences" on an *absolute scale* (i.e., a ratio scale) is likely to be unsuccessful. The method we have employed, however, does provide interval scales regarding the psychological dimensions of interest.

The "interval scale" property of the scales we have derived does depend on the assumption that the discriminal dispersions of all the stimuli (scenarios) are approximately equal. We cannot verify that this is so from the data, but we feel the assumption is warranted, in that the resulting scaling model provides a useful description of the judgmental data we have obtained. Actually, we know that another assumption implicit in the Thurstone technique, that the "distances" between scenarios can be represented *on a single* dimension, is definitely violated. This violation is not to be regarded as a "fault" of the data; it reflects the interesting discovery that judgments of the scenarios with respect to the important tactical decision variables are inherently multidimensional. This is discussed under the section "Multidimensional Scaling." Nonetheless, we feel that the unidimensional scales derived by the Thurstone technique are of considerable practical value in displaying the results of the survey.

We now turn to a description of the practical application of the Thurstone scaling technique to the scenario ranking data. When a judge has rank-ordered the scenarios, it is assumed that *the order of any pair of those scenarios fairly represents his judgment of them at the time he performed the ranking.* Thus, each rank-ordering of the 17 tactical scenarios by a subject is assumed to represent his comparative judgment regarding the 17x16/2 = 136 possible pairs of scenarios with respect to the decision variable of interest. (In the "paired comparison" technique which is sometimes used, each of these 136 pairs would *individually* be presented to the subject; however, the method of rank ordering, which implicitly produces the same information, is much faster to administer.) Thus, when N judges have rank-ordered the scenarios according to a given set of instructions, the experimenter can examine the N judgments of each of the 136 pairs of scenarios, and form the proportion "p_{ij}" referred to earlier, representing the proportion of the N subjects who judged scenario "j" to be greater than scenario "i." These data are conveniently represented in a *proportion matrix*, P, shown in Figure 1. Note that the diagonal elements, p_{ii}, representing the comparison of a scenario with itself, are assigned the proportion 0.5. Also note that symmetric elements must sum to unity

($p_{ij} + p_{ji} = 1$), so half the elements are redundant.

We now turn to the "distance" estimates. We have seen that $\Phi^{-1}[p_{ij}]$ is an estimate of the distance z_{ij} between scenarios "i" and "j" (in units of the standard deviation of the discriminal difference). Therefore, we perform the transformation (i.e., determine the unit normal deviates corresponding to the observed proportions) to obtain the matrix Z shown in Figure 2. Because of the properties of the P matrix, the diagonal elements of the Z matrix are all zero, and the matrix is skew-symmetric (i.e., $z_{ij} = -z_{ji}$).

We know from the theoretical discussion of the Thurstone scaling model that the element z_{ij} of the Z matrix is an estimate of the "distance" between scenarios "i" and "j." Therefore, we could place the scenarios on a scale using the following technique. First, assign scenario 1 an arbitrary location on the scale. Then, place scenario 2 a distance $z_{1,2}$ from scenario 1 on the scale; then place scenario 3 a distance $z_{2,3}$ from scenario 2; place scenario 4 a distance $z_{3,4}$ from scenario 3; and so on. The resulting scale would graphically represent (within a linear transformation) an estimate of the scenario positions on the psychological dimension of interest.

However, this technique makes poor use of the experimental data, simply because not all are used; those elements z_{ij} for which $j \neq i+1$ would not be employed. In the above technique, only *one estimate* of the distance between scenarios "i" and "j" is made, that represented by the element z_{ij}. Yet, other estimates are possible, because *we can consider the difference between the distance estimates from scenario "i" and scenario "j" to some third scenario, "k," to also be an estimate of the distance between scenarios "i" and "j."* That is, $z_{ik} - z_{jk} = z'_{ij}$, where the "prime" indicates an estimate of z_{ij}. For the present case, there are 17 such estimates possible,

$$\begin{array}{c|c|c|c|c|c|c|}
 & 1 & 2 & 3 & & 16 & 17 \\ \hline
1 & 0.5 & p_{1,2} & p_{1,3} & \cdots & p_{1,16} & p_{1,17} \\ \hline
2 & p_{2,1} & 0.5 & p_{2,3} & \cdots & p_{2,16} & p_{2,17} \\ \hline
3 & p_{3,1} & p_{3,2} & 0.5 & \cdots & p_{3,16} & p_{3,17} \\ \hline
 & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \hline
16 & p_{16,1} & p_{16,2} & p_{16,3} & \cdots & 0.5 & p_{16,17} \\ \hline
17 & p_{17,1} & p_{17,2} & p_{17,3} & & p_{17,16} & 0.5 \\ \hline
\end{array}$$

Figure 1. A proportion matrix.

$$\begin{array}{c|c|c|c|c|c|c|}
 & 1 & 2 & 3 & & 16 & 17 \\ \hline
1 & 0 & z_{1,2} & z_{1,3} & \cdots & z_{1,16} & z_{1,17} \\ \hline
2 & z_{2,1} & 0 & z_{2,3} & \cdots & z_{2,16} & z_{2,17} \\ \hline
3 & z_{3,1} & z_{3,2} & 0 & \cdots & z_{3,16} & z_{3,17} \\ \hline
 & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \hline
16 & z_{16,1} & z_{16,2} & z_{16,3} & \cdots & 0 & z_{16,17} \\ \hline
17 & z_{17,1} & z_{17,2} & z_{17,3} & & z_{17,16} & 0 \\ \hline \hline
\frac{1}{17}\sum_{j} z_{ij} & z_1^* & z_2^* & z_3^* & & z_{16}^* & z_{17}^* \\ \hline
\end{array}$$

Figure 2. A Z-matrix.

one for each value of "k," the "reference scenario." (When $k=j$, we have $z_{ij} - z_{jj} = z_{ij} - 0 = z_{ij}$, the *only* estimate used in the elementary scaling procedure suggested above; and when $k=i$, we have $z_{ii} - z_{ji} = 0 - z_{ji} = -z_{ji}$, which is identical to z_{ij}.)

Thus, the experimental data contain 17 estimates of the distances between scenarios 1 and 2; between 2 and 3; and so on. *Because of experimental error, not all these estimates for a given scenario pair will be numerically equal.* To obtain the "best" estimate of the distance between a given pair of scenarios, the *arithmetic average of the 17 estimates at hand* naturally suggests itself. Mosteller (1951) has shown that doing this will provide a least-squares estimate of the "true" z_{ij}. Thus, the "best" estimate, z^*_{ij}, is given by:

$$z^*_{ij} = \frac{1}{17} \sum_{k=1}^{17} (z_{ik} - z_{jk})$$

But, note that this is equivalent to:

$$z^*_{ij} = \frac{1}{17} \sum_{k=1}^{17} z_{ik} - \frac{1}{17} \sum_{k=1}^{17} z_{jk}$$

which is just the difference between the mean values of the cells in columns "i" and "j" of the Z matrix. These column means are shown in the bottom row of the Z matrix illustrated in Figure 2, denoted by "z^*_i."

Thus, if each scenario, "i," *is placed on a scale at the value of its column mean,* "z^*_i," *the distances between the scenarios in the resulting graphical presentation constitute the "best" representation of the scenario positions on the psychological dimension of interest,* within a linear transformation. That is, the distance relationships are represented, but the origin and units of the psychological continuum are not defined. The origin of the graphical representation will be at the mean of the scale values of the 17 scenarios (as an incidental result of the scaling procedure), and the unit will be the standard deviation of the discriminal difference.

In the "Results" sections following, the scenario scale positions are shown graphically as derived by this technique.

Results: Sonar Contact Reporting Thresholds

The scale values derived for the scenarios on the "Sonar Contact Reporting" continuum from the entire survey sample are shown in Figure 3. At the top, corresponding to the "freest" or most "aggressive" reporting criterion (corresponding to the least strict value of likelihood-ratio criterion) we find the following scenarios:

3. Wartime screening for merchant convoy
11. Wartime screening for amphibious task group
9. Wartime strike group operations
16. Wartime hunter-killer operations

The property which characterizes these scenarios is obvious: they are wartime situations.

The next group of scenarios include the following:

2. Unidentified contact, datum 1 hour old
10. Unidentified contact, datum 4 hours old
13. Condition III steaming in the Mediterranean

These scenarios have in common the possibility of contact with Soviet submarines in international waters.

The next group of scenarios include the following:

7. Opposed sortie ASW exercise
14. High valued target, ASW screening exercise
6. Annual ASW competitive exercise
15. ASW type training
4. ASW refresher training

These scenarios all represent ASW exercise or training situations.

The last distinct group of scenarios on the "Sonar Contact Reporting" continuum, in the direction of a "strict" or "conservative" contact reporting criterion (corresponding to a

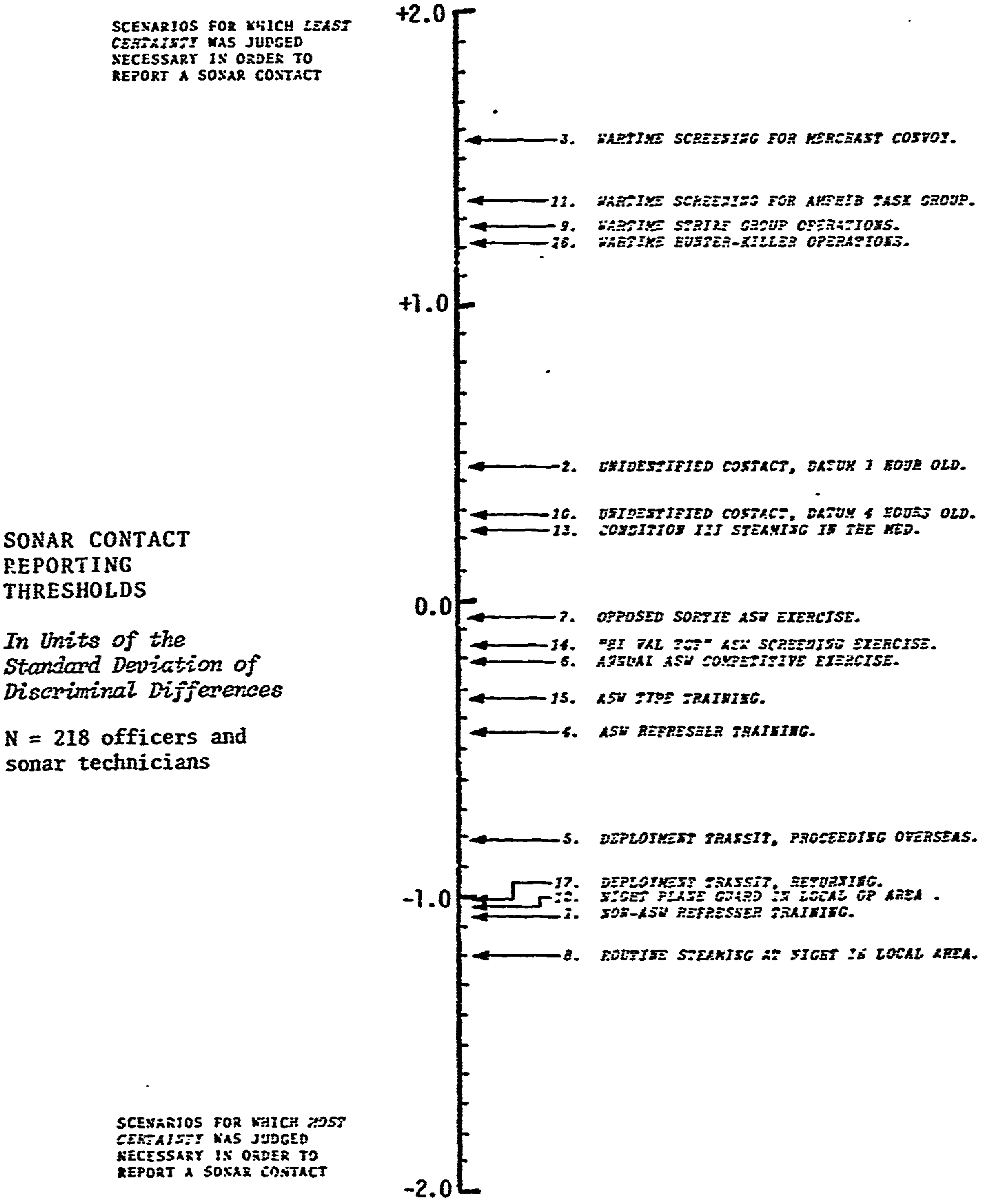


Figure 3. Sonar contact reporting thresholds for 17 tactical scenarios.

stringent value of likelihood-ratio criterion), include the following scenarios:

5. Deployment transit, proceeding overseas
17. Deployment transit, returning
12. Night plane guard in local op area
1. Non-ASW refresher training
8. Routine steaming at night in a local area

Each of these scenarios involves a routine peacetime task not primarily connected with ASW.

Thus, on the psychological continuum of sonar contact reporting, we have a complete spectrum of reporting criteria represented, from Scenario 8, "routine steaming at night in a local area," during which considerable contact classification certainty is felt necessary to cause sonar to arouse and alert the ship; to Scenario 3, "escorting a merchant convoy in the North Atlantic during a hot-war situation," in which very little evidence is felt necessary to cause the sonar operator to report. The positions of the scenarios on this axis seem easily understood, and are in accord with our expectations.

The total sample of respondents was broken down in several ways in an attempt to identify significant differences between groups. Officers' scale values were compared to sonar operators'; Atlantic Fleet respondents were compared to Pacific Fleet respondents; "junior" (defined as less than four years' service for enlisted personnel and less than eight years' service for officers) were compared to "senior" personnel; and commanding officers' scale values were compared to all others. No significant differences among the various groups were noted. However, there *is* a breakdown which identifies significantly different groups along this continuum; this breakdown will be discussed under "False Contact Consequences."

Results: Missed Contact Consequences

The scale values for the scenarios on the "Missed Contact Consequences" continuum as derived from the total survey sample are shown in Figure 4. (No differences among groups were noted.) The order of the scenarios along this scale is substantially the same as that along the "Sonar Contact Reporting" scale, but *the wartime scenarios are significantly further removed* from the others, in the direction of more severe consequences. This is not surprising. The contact reporting threshold is dependent on all the other decision variables, and the relatively low wartime *a priori* probability of contact, compared to the ASW exercise situations, probably acts to prevent an extreme separation of the war scenarios from the others on the "Sonar Contact Reporting Threshold" scale. However, the "Missed Contact Consequences" scale only involves a single variable of the reporting decision: consequences of the error of omission. It seems easy to understand why, on this scale, the scenarios can be classified dichotomously: war and peace.

Results: False Contact Consequences

In marked contrast to the other scales, the results for the total sample in the scale of False Contact Consequences showed little meaningful separation of the scenarios (this is not illustrated). None of the previously mentioned breakdowns (i.e., officer/enlisted, Atlantic/Pacific, junior/senior, CO/other) results in any significant separation of the scenarios, either. However, by direct examination of individual answer sheets, it was observed that the respondents fell into three categories which were previously unsuspected. We shall refer to these as "Groups I, II, and III." When the total sample is broken into Groups I, II, and III, the scaling results were as shown in Figures 5, 6, and 7. Group I, composed of 51 officers and 68 operators (approximately 55% of the total sample), ranked *the wartime scenarios* as having the *most severe consequences from false contacts*. The resulting scale for Group I is not greatly different from the Missed Contact Consequences scale for the entire sample. (See Figure 5.)

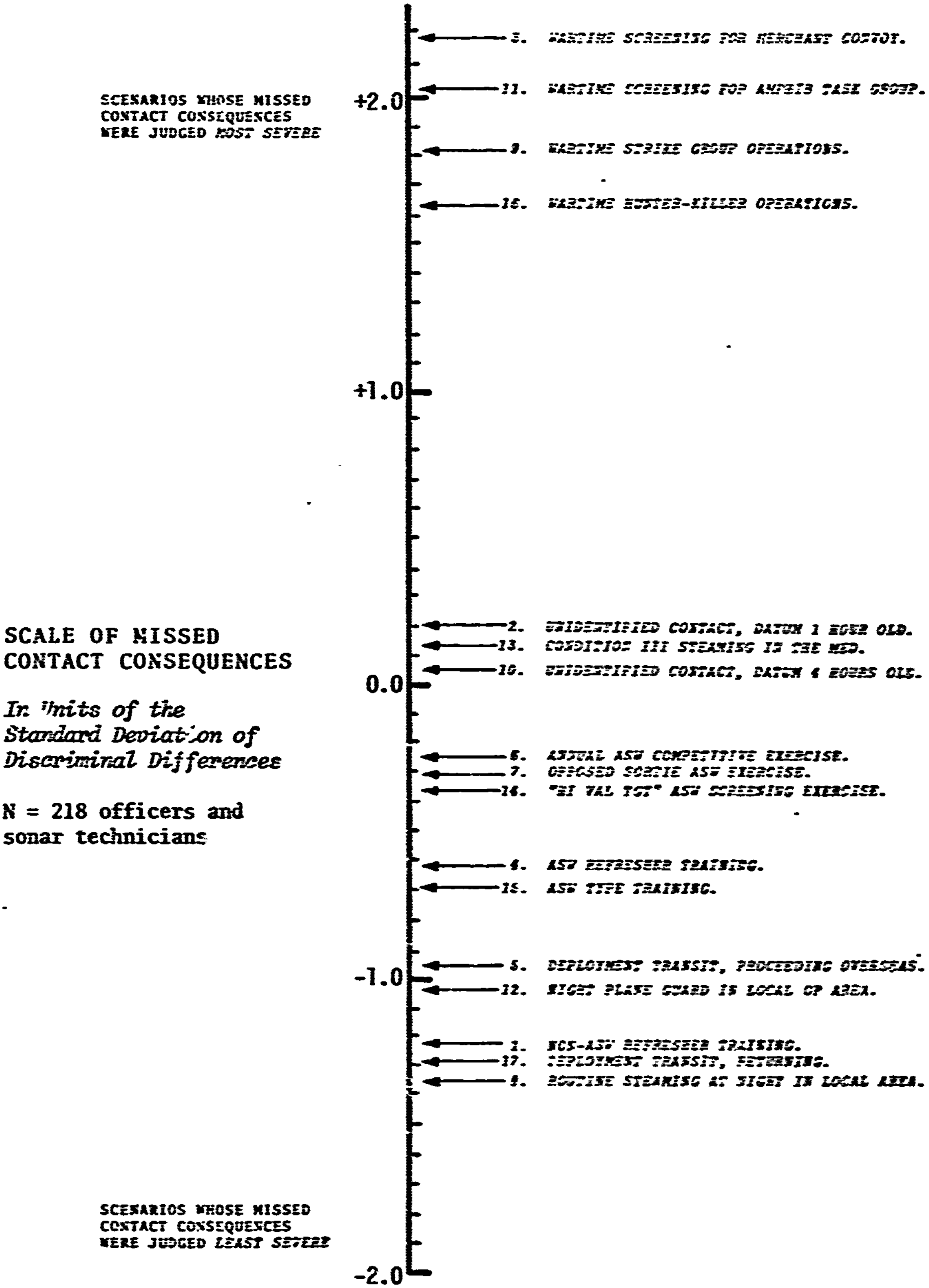


Figure 4. Position of 17 tactical scenarios on the scale of missed contact consequences.

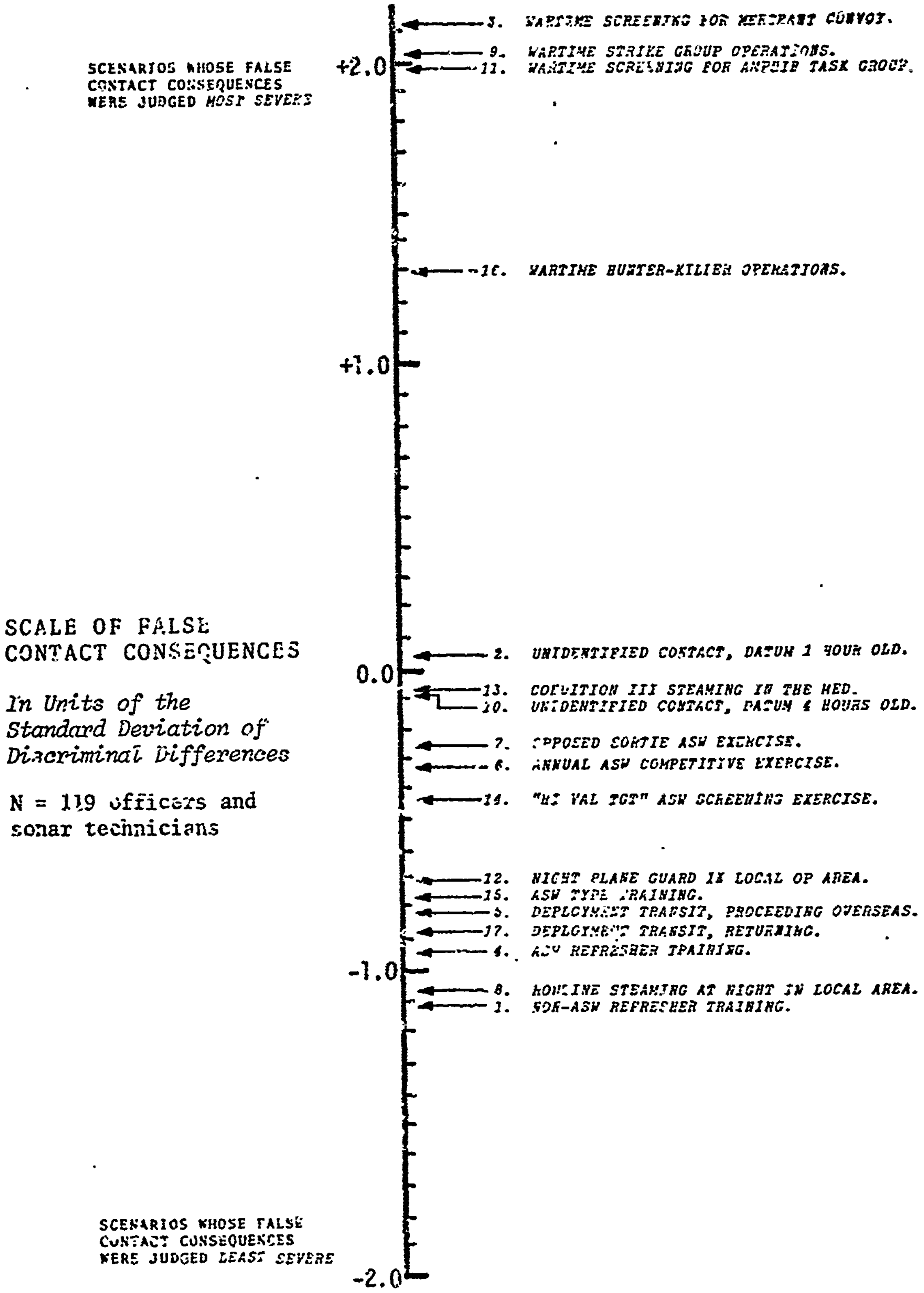


Figure 5. Positions of 17 tactical scenarios on the scale of false contact consequences *as judged by "Group I" subjects.*

Group II, composed of 26 officers and 34 operators (about 28% of the total sample), ranked the scenarios oppositely: they ranked the wartime scenarios as having the *least severe consequences.* The scale values for Group II are nearly mirror images of those for Group I; the Pearson product-moment correlation coefficient between the scale values for Groups I and II is -0.97. (See Figure 6.)

The members of Group III are characterized by unpredictability in their placement of the scenarios in rank order; that is, they are likely to have placed one wartime situation toward one end of the scale, but another wartime scenario toward the other end. The rank of a *given* scenario, however, was not consistent among the judges in Group III. This, of course, indicates that the members of Group III did not distinguish the scenarios with respect to false contact consequences. (See Figure 7.) This group consisted of 19 officers and 18 operators (approximately 17% of the sample).

After the breakdown of the sample according to attitudes regarding false contact consequences was discovered, a recomputation of scale values for the other dimensions was performed separately for Groups I, II, and III. This reanalysis showed that false contact con ·quence attitudes had no significant effect on the scale values for Missed Contact Consequences, Subjective Probability of Contact, or Delayed Contact Reporting Consequences. This indicates that the other decision variables were independent of attitudes toward false contact consequences. However, false contact consequences would certainly be expected to be related to the "degree of confidence" felt necessary for reporting sonar contacts; that is, to the setting of the reporting criterion. This is predicted by decision theory, and certainly is intuitively compelling as well. And, indeed, when scale values on the "Sonar Contact Reporting Threshold" continuum were recomputed separately for Groups I and II, a significant difference was

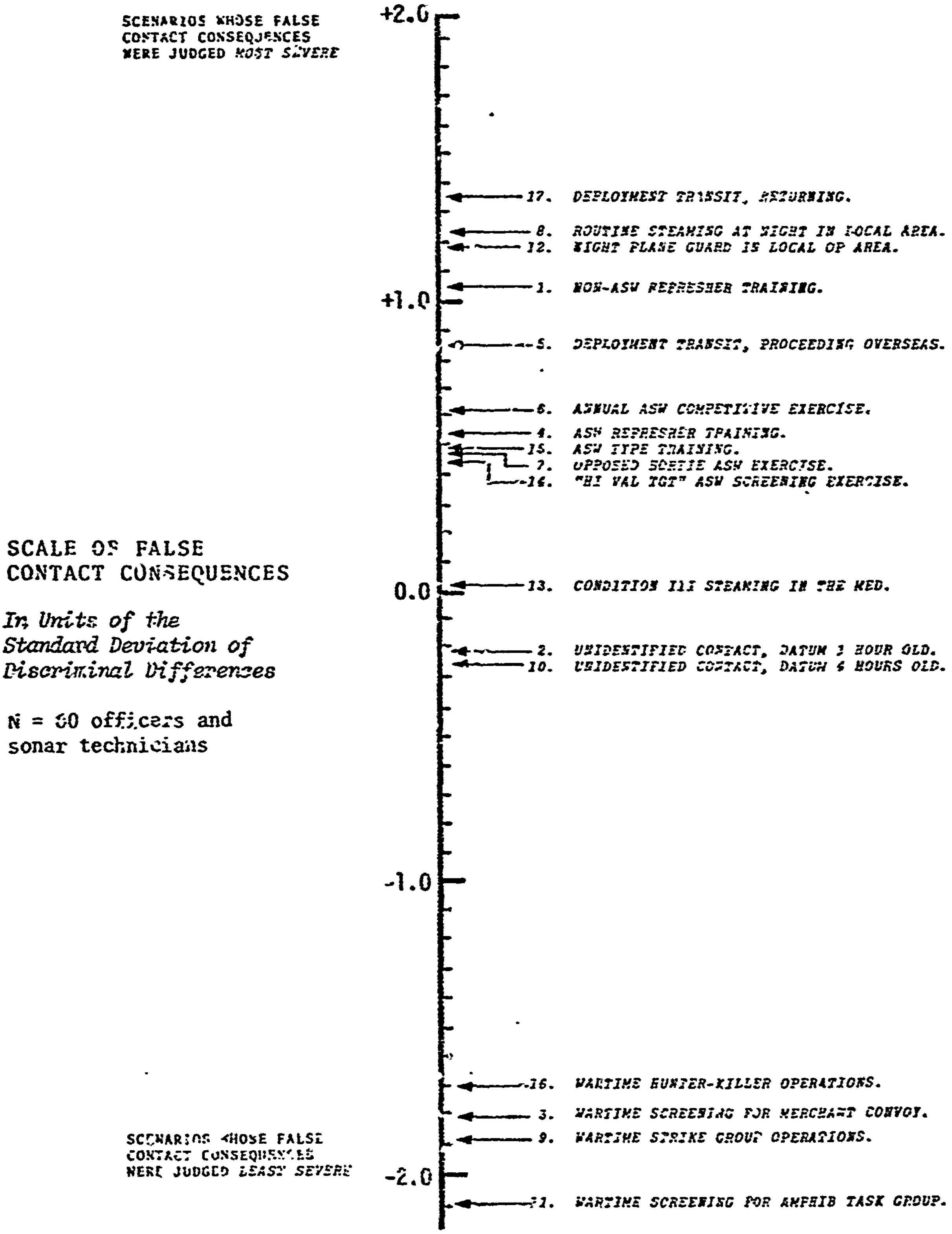


Figure 6. Positions of 17 tactical scenarios on the scale of false contact consequences *as judged by "Group II" subjects.*

]

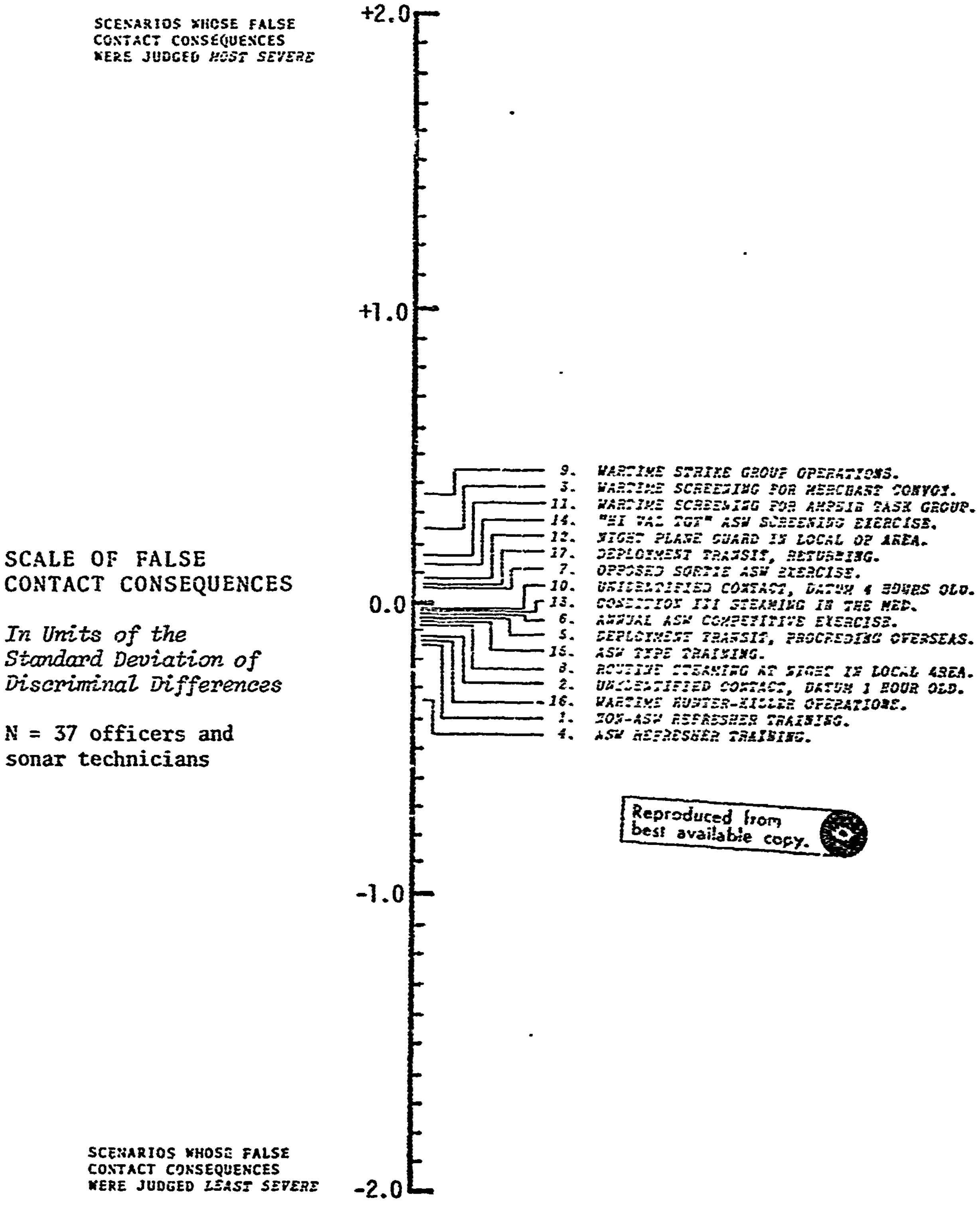


Figure 7. Positions of 17 tactical scenarios on the scale of false contact consequences *as judged by "Group III" subjects.*

noted, as can be seen in Figure 8. Since severe false contact consequences should induce a *more cautious* reporting criterion, we would expect the wartime scenarios for Group I to be found *farther down* on the scale than those of Group II, since Group I perceived false contact consequences to be worst during war. This effect is seen in the scaling results. Likewise, we would expect the routine peacetime scenarios for Group II to be *farther down* on the scale than those of Group I, since Group II perceived false contact consequences to be worst during *these* scenarios. This effect is also seen. These effects are particularly interesting since all the respondents performed the ranking on the sonar contact reporting continuum *prior* to ranking according to false contact consequences, or any of the other decision variables. Thus, it appears that, without their being explicitly called to the respondent's attention, false contact consequences were weighed, consciously or unconsciously, in performing the ranking according to the "sonar contact reporting" instructions. This observation supports the expectation that the subjective values of these variables are influential in actual sonar contact reporting behavior. By way of interpretation, it would seem that members of Group II would be prone to a much greater shift in willingness to prosecute an unknown contact in the eventuality of war than members of Group I.

Analysis of the survey data showed no way to predict the membership of Groups I, II, or III on the basis of the personnel data collected. No group had a membership which was characteristically officer, enlisted, Atlantic Fleet, Pacific Fleet, more experienced, less experienced, or composed of commanding officers. It follows that whatever influences the attitudes of Groups I, II, and III have on decision-making behavior in ASW are now distributed unpredictably. As we have seen, however, it is possible to identify these attitudes by means of the survey technique; and it is probable

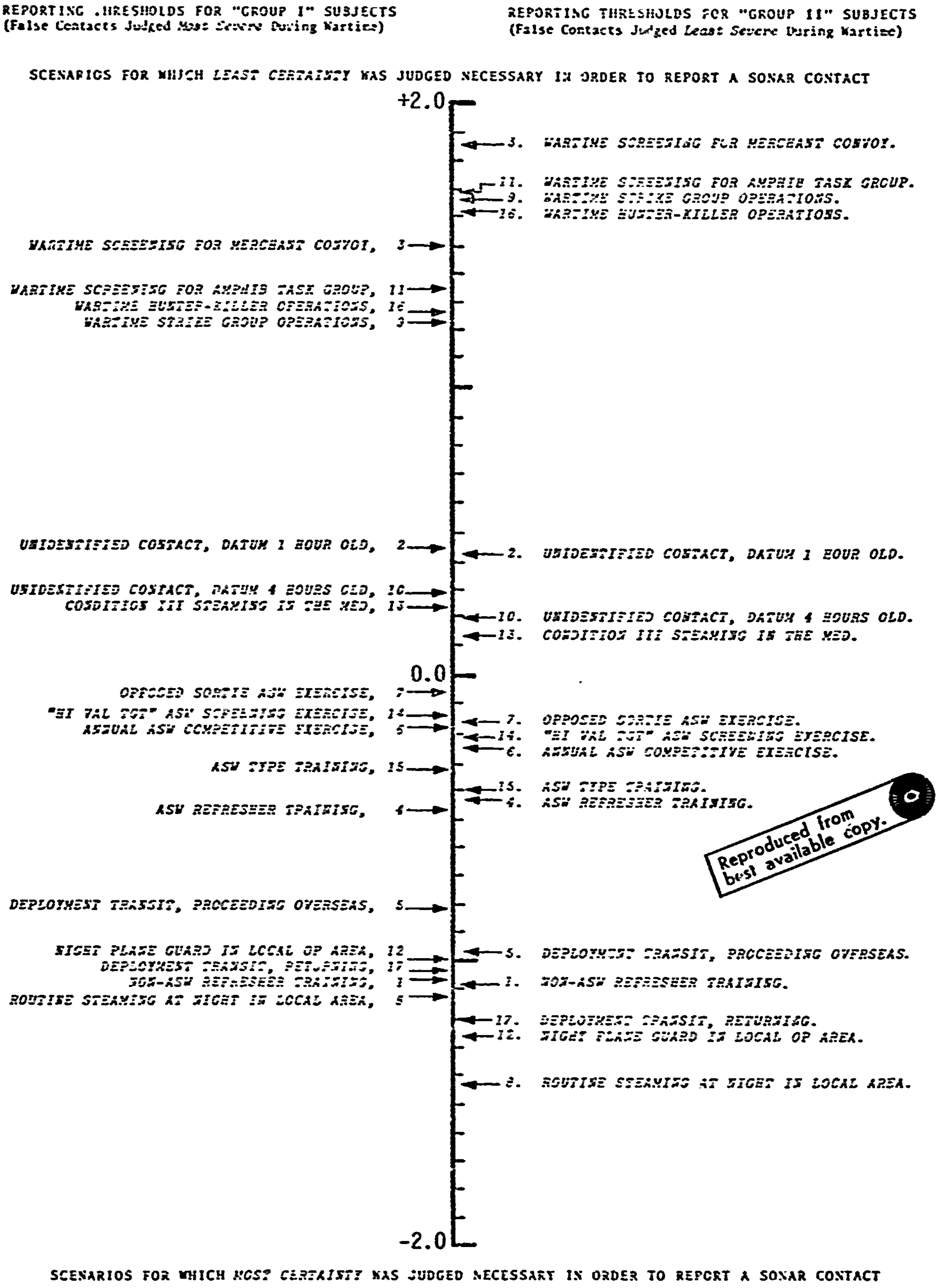


Figure 8. Sonar contact reporting thresholds for two groups separated by attitudes toward false contacts.

that these views can be modified through training if it is found to be desirable. However, the question of the relationship of these *attitudes* to actual *decision-making behavior* first must be answered. This question will be considered in the "Experiment" section of this report.

Results: Subjective Probability of Submarine Contact

The scale values of the scenarios along the "Subjective Probability of Making Contact" scale derived from the total sample are shown in Figure 9. There were no differences in scale values assigned among any of the identified subgroups. It is obvious that these scale values reflect a different property of the tactical scenarios than the other dimensions scaled in this survey. This is not surprising since, among the decision variables, prior probability stands apart; the other variables involve "consequence" considerations which we expect to be interrelated, but which have little to do with "subjective likelihood of making contact." The latter is determined principally by the expected deployment pattern of the "target" submarines. Both theoretically and practically, the subjective probability dimension can be expected to be orthogonal to the "consequence" dimensions.

The scenarios fall into two distinct groups: the first including the exercise, the wartime, and the unidentified contact scenarios; and the second, in the direction of least likelihood of contact, including the peacetime non-ASW scenarios. It is very obvious that expectation of contact during peacetime non-ASW operations is extremely low. This general grouping of the scenarios was expected; but it is of interest that the wartime scenarios were placed as high on the scale as they were, at the same average position as the "unidentified contact" scenarios, and quite near the ASW exercise situations.

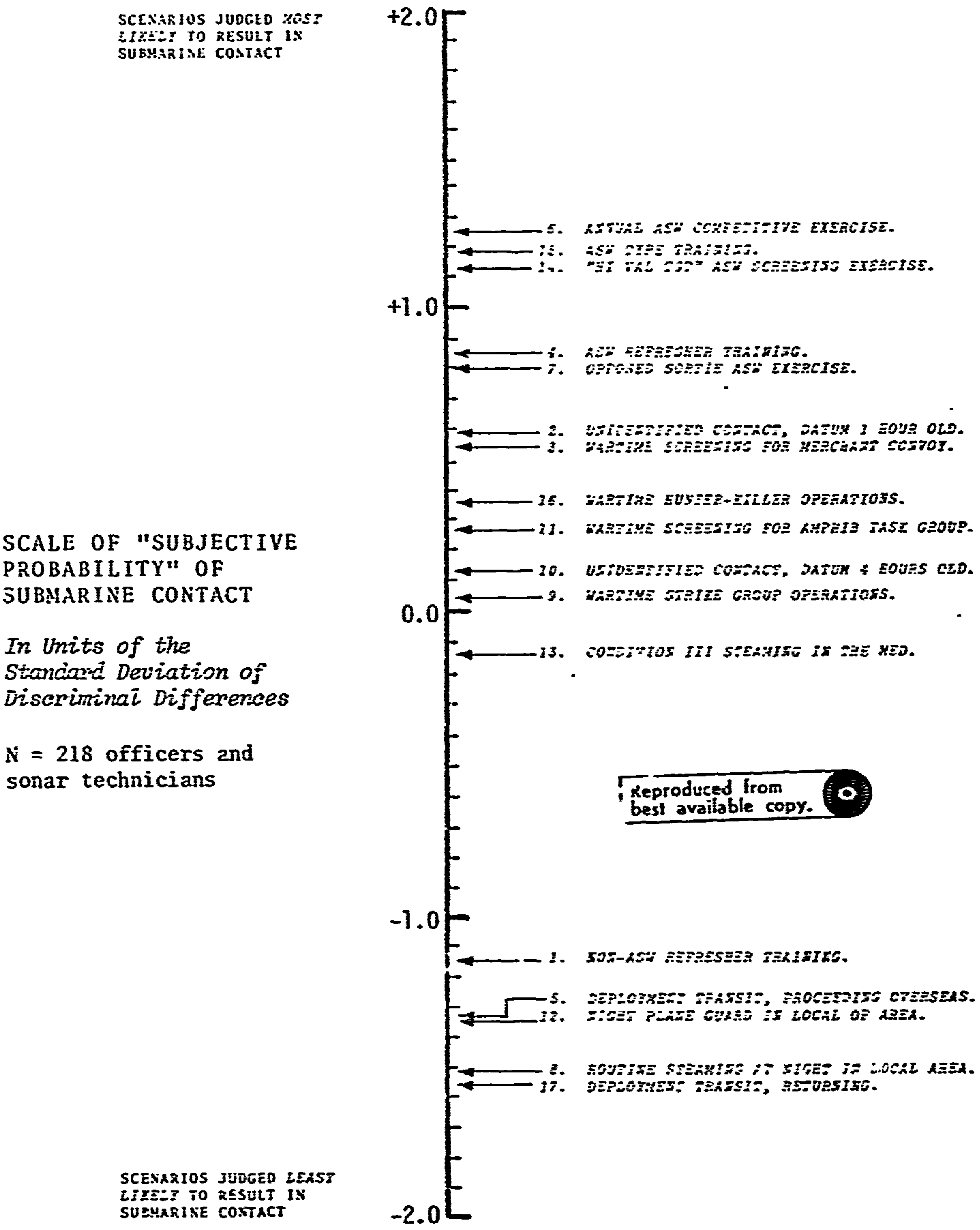


Figure 9. Positions of 17 tactical scenarios on the scale of "subjective prcbability" of submarine contact.

Results: Consequences of Delayed Contact Reporting

Figure 10 shows the results of scaling the rank orderings according to the "response time" set of instructions for the total survey sample. It can be seen that the consequences of delayed contact reporting were evaluated in very much the same way as the consequences of missed contacts, shown in Figure 4. On the latter scale, there is a more distinct differentiation of the "cold-war" scenarios from the "exercise" scenarios, and a greater separation of the "wartime" scenarios from the others, which seems reasonable. Nonetheless, it is evident that the "consequence" considerations involved in missed contacts, false contacts (at least according to the "Group I" attitude), and delayed reporting are very strongly related, and that these considerations make the principal contribution to the character of the "sonar contact reporting threshold" scale. The "threshold" scale differs from the "consequence" scales in showing a less extreme separation of the wartime scenarios from the others, which can be attributed to the influence of the "subjective probability" variable, which tends to produce a "conservative" shift in the reporting criterion for the (relatively) less probable wartime contacts, and an "aggressive" shift for the (relatively) probable exercise and training sonar contacts. The separate scaling of the "reporting threshold" data by Group I/II "false contact consequence" attitudes (Figure 8) also reveals a significant quantitative influence upon the sonar contact reporting thresholds, but no gross qualitative effect (such as an inversion of the scale corresponding to the inverted Group II false contact consequence scale).

Thus, it may be concluded that the situational variable of greatest importance in determining the sonar contact reporting threshold is that which gives rise to the perceived consequences of delayed sonar contact reporting, and of missed contacts; a situational variable we may aptly call the "threat."

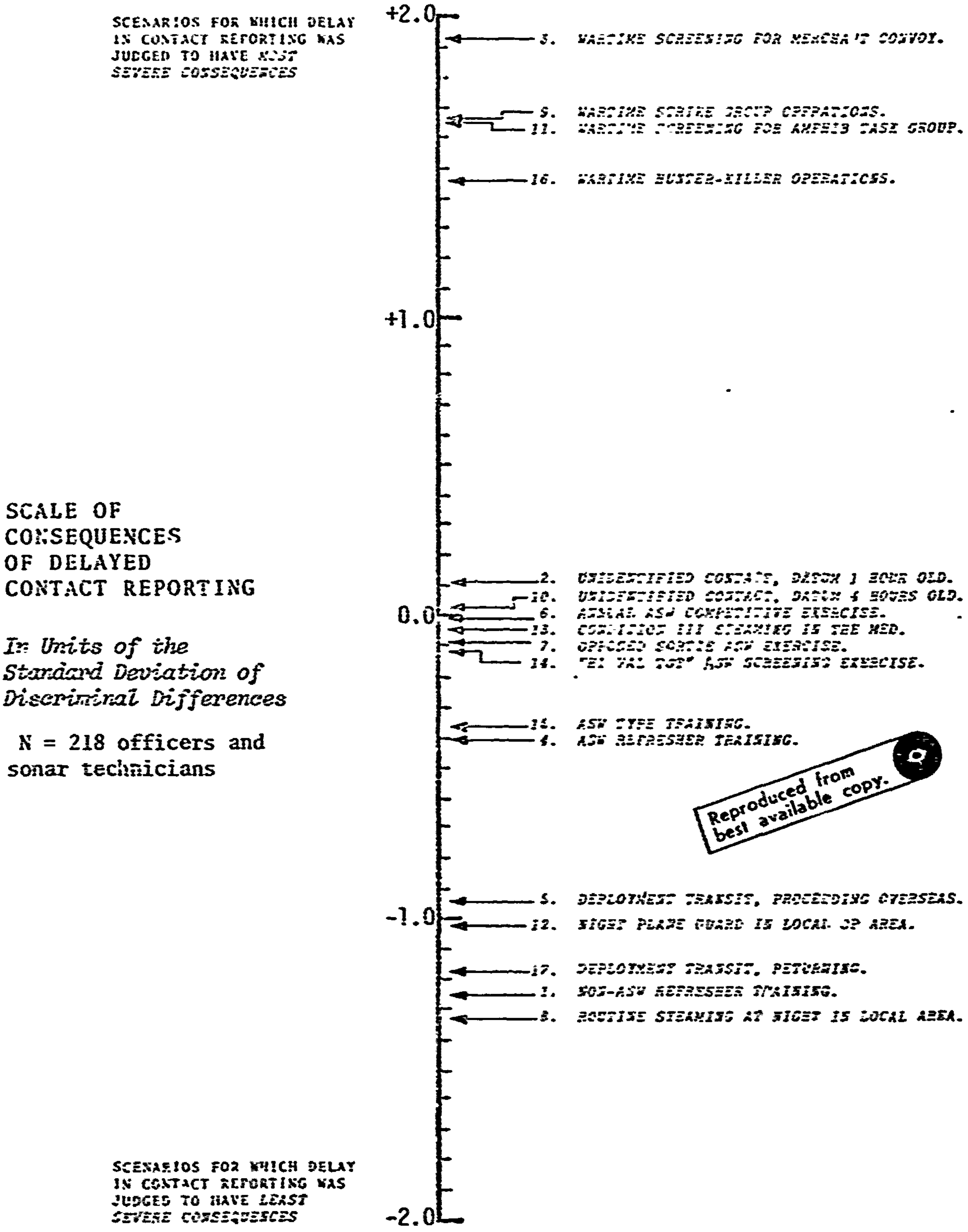


Figure 10. Positions of 17 tactical scenarios on the scale of consequences of delayed contact reporting.

Multidimensional Scaling

In the Thurstone scaling technique it is assumed that the psychological "distances" between stimuli may be adequately represented in a single dimension; that is, they may be presented on a single scale, as we have done. However, if the aspect of the stimuli being regarded by the judges is inherently multidimensional, the distances between stimuli cannot be truly represented on a one-dimensional scale. Most often, multidimensional stimuli are regarded as being representable by points in Euclidean space, forming a simple geometrical model which (at least up to 3 dimensions) is analogous to the familiar spatial surroundings in which we perceive ourselves. Then, "distance" between points has its familiar meaning.

Since the ranking task forces each judge to place the stimuli in a one-dimensional ordinal arrangement, it is not immediately evident that multidimensional information can be recovered from rank-order data. Nonetheless, it can, because of the variability among judges concerning the precise way in which each forces the stimuli onto a single dimension. The same basic assumption that was made for the Thurstone technique is made for multidimensional scaling: that the *relationship of proportion of agreement among judges concerning two stimuli to the psychological distance between the stimuli is monotonic.*

This was the *first* assumption made for the Thurstone technique; it was necessary to make another, more restrictive, assumption (that there is a "discriminal process" with a Gaussian distribution) to obtain actual scale values. However, we do not need to make the latter assumption for the multidimensional scaling technique that we employed, which is an advantage of the technique. We need only the assumption that the judged "dissimilarities" between stimuli (measured by the proportion of agreement among the judges) is monotonically related to the "distances" between stimuli in the psychological space, so that the rank order of the "dissimilarities" for all pairs of stimuli is identical with the rank order of the "distances" for all pairs of stimuli. If this is so, then a spatial arrangement of the stimuli which corresponds to the original configuration of the stimulus positions in psychological space *may be recovered from the rank order of the dissimilarities* (except for a possible rotation, translation, or expansion, which does not affect the rank order of the distances). To grasp this intuitively, consider a collection of many objects (for example, marbles) to be strewn about in Euclidean two-space (for example, the floor of a handball court). If there are lots of marbles, you can readily appreciate the fact that a *given* marble can't be moved very far in any direction without disturbing the rank order of the distances between that marble and all the others, since, as it is moved closer to some, it will be moved farther from others. Thus, the specification of the rank order of the distances from the given marble to all the others pretty well fixes its position on the floor of the court. And, since the "dissimilarities" are assumed to be in the same rank order as the distances, the rank order of the dissimilarities *also* fixes the position of the given marble.

You can also appreciate the practical difficulty of trying to arrange marbles in a handball court to satisfy a long list of rank-ordered distances or dissimilarities which should exist among them. This very cumbersome, iterative reconstruction, however, can be handled expeditiously in the digital computer. The principles and practical implementation of this technique were first given by Shepard (1962a; 1962b), and refined by Kruskal (1964a; 1964b); and we have used a computer program (TORSCA) prepared by Young and Torgerson (1967) to accomplish the necessary manipulations.

The "dissimilarities" used by the program were derived from the proportion matrix (Figure 1), and the program arranged the stimuli (scenarios) in two dimensions so that the distances between the scenarios were (as closely as possible) in the same rank order as the judged dissimilarities between the scenarios. (We also scaled the data in higher dimensions, and concluded that the two-dimensional solution is most appropriate.)

The configurations of the scenarios derived from the rankings according to missed contact consequences and consequences of delayed contact reporting are shown in Figures 11 and 12, respectively. The wartime scenarios (3, 9, 11, and 16) are omitted from these figures, because their extreme separation from the peacetime scenarios masks the interesting relationships among the latter, if the computer program is forced to include them.

It is evident from these figures that a second dimension is needed to completely represent the judgmental data. It is also evident that the horizontal dimension, labeled "tactical consequence of error," corresponds most strongly to the single dimension upon which the scenarios were scaled by the Thurstone technique (Figures 4 and 10), that is, the projections of the scenarios onto the horizontal dimensions of Figures 11 and 12 lead to representations very similar to those of Figures 4 and 10.

What is represented by the second dimension (the vertical axis)? We postulate that this dimension is related to "personal consequence" of error. In the case of the ASW exercise and training scenarios, for example, we see that the "exercise" situations are distinctly separated from the "training" situations on the vertical dimension. The ASW exercise situations have an "evaluative" connotation which may make errors of greater "personal" consequence than in the training situations,

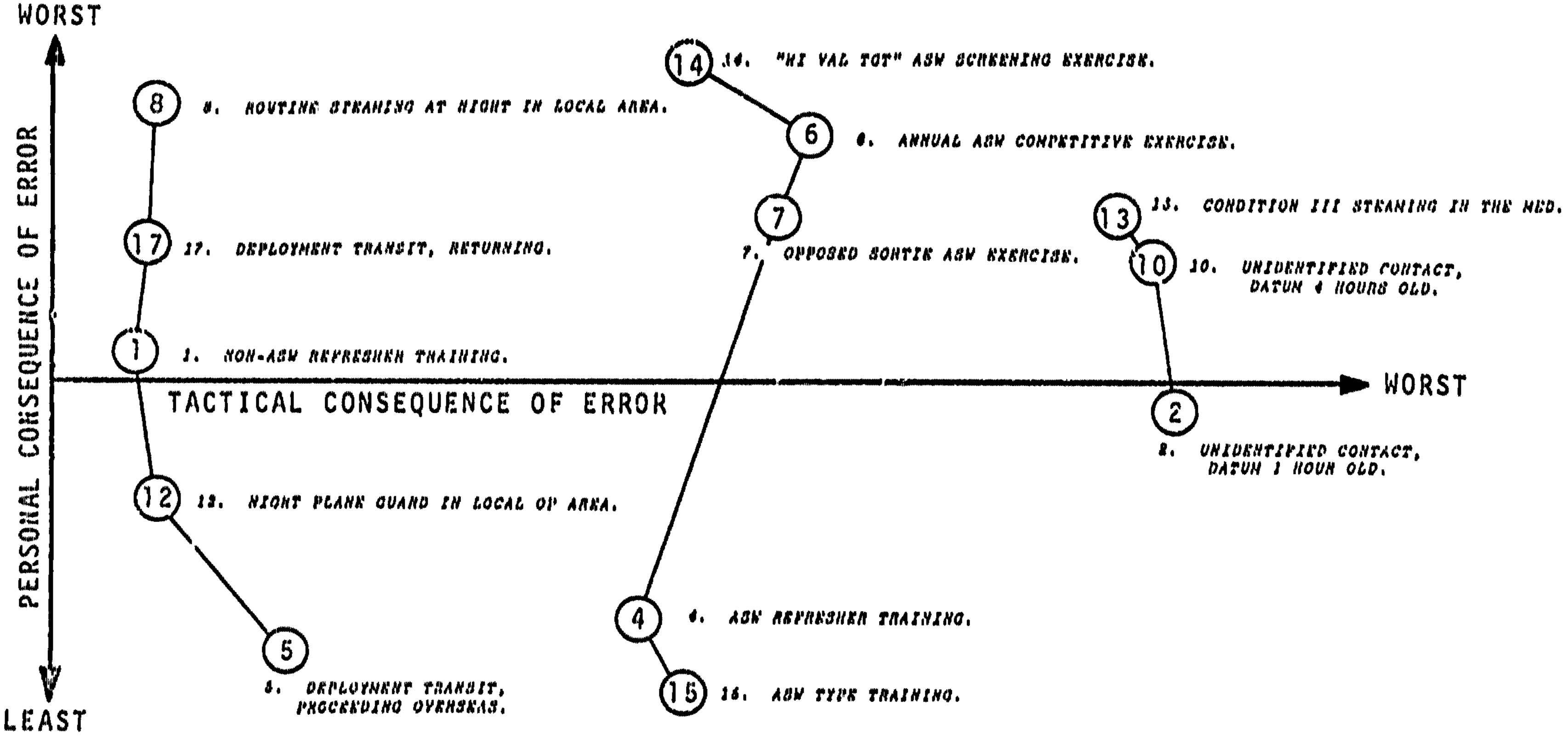


Figure 17. Responses concerning "missed contact consequences" scaled in two dimensions.

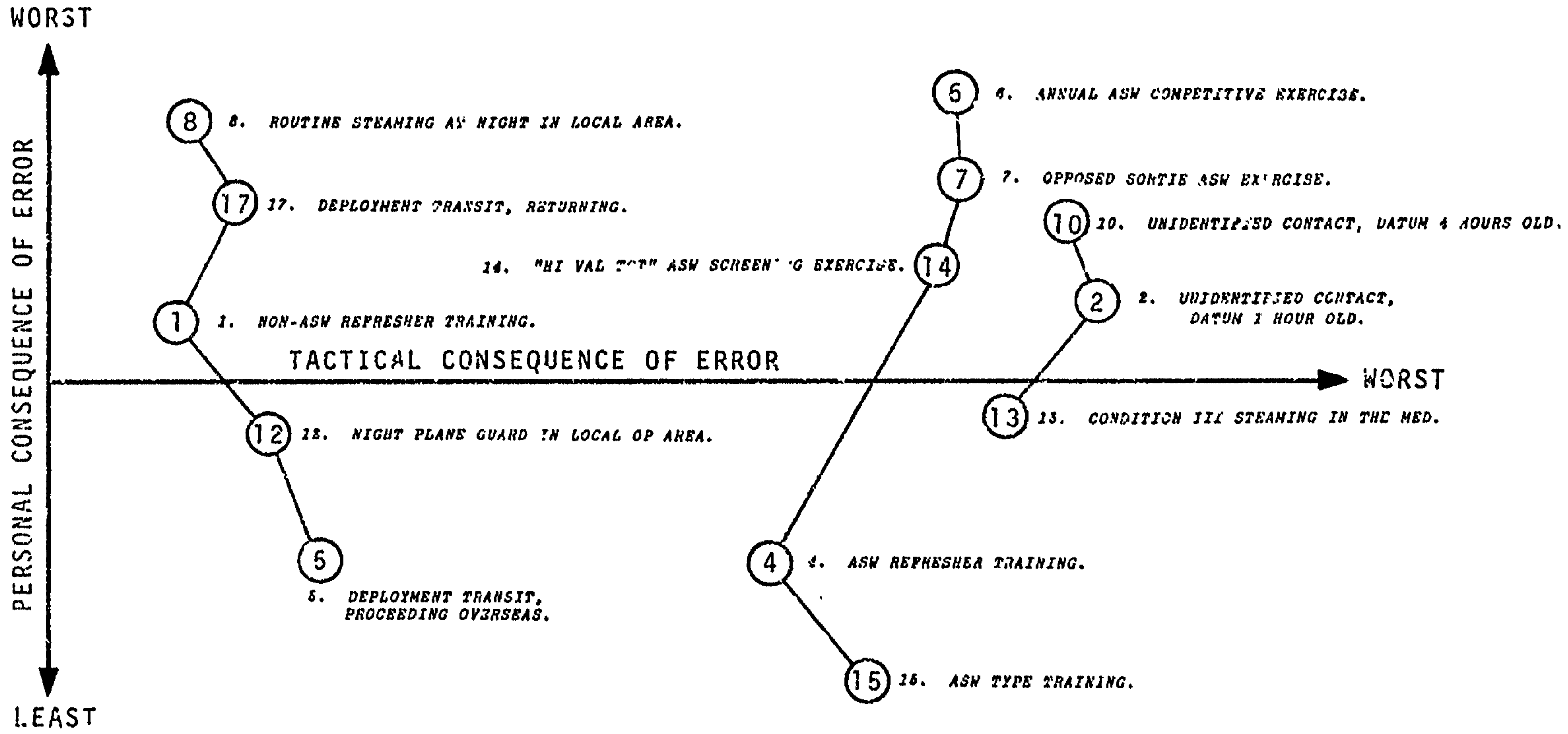


Figure 12. Responses concerning "consequences of delayed contact reporting" scaled in two dimensions.

which are by their nature more tolerant of error. Thus, perceived "personal consequence" seems to increase as we move upward on the vertical axis.

The non-ASW situations appear to be ordered along the vertical dimension by what we will call "general arousal level" of the ship, decreasing as we go up the scale. In the "deployment transit, proceeding overseas" there would be a general, high level of arousal throughout the ship; diminishing, as we consider the scenarios in their order along the scale, to "routine steaming at night," in which most of the ship's company is asleep. We suggest that an "error" on the part of the sonar operator is perceived to be a potentially smaller perturbation of the "status quo" when the ship is in a "high" arousal state, and therefore of lesser personal consequence, than an error committed during a state of *lower* general arousal (which, in the extreme, might involve waking a sleeping ship's company, as a consequence of "over-eager" contact reporting; such behavior might be commended by the conscientious commanding officer, but we are led to believe that the typical ship's company is not so likely to offer positive reinforcement of this behavior).

The "cold war" scenarios (2, 10, and 13) are not consistently ordered on the "personal consequence" dimensions between Figures 11 and 12. We attribute this to experimental error, owing to their relatively small separation on the vertical dimension. A similar inconsistency is seen for scenarios 14, 6, and 7 between Figures 11 and 12; these scenarios are also very close together.

The discovery of the "personal consequence" dimension does not significantly alter any of the conclusions drawn from the scaling via the Thurstone technique. The effect of the separation of the scenarios in this *second* psychological dimension, when the scenarios are forced onto a *single* scale by the

Thurstone technique, is to introduce a little more separation among the scenarios of a group (i.e., non-ASW; exercise; "cold war," wartime) than the *projections* of the two-dimensional scenario positions onto the "tactical consequence" axis would show. Thus, the scenarios are somewhat more tightly "grouped" on the psychological "tactical consequence" scale than the Thurstone scaling indicates. The unidimensional representation remains a useful one, but it is interesting that "personal consequences" seem to play an identifiable role in the judgment of decision consequences.

Questionnaire Results

In addition to the ranking tasks, each survey respondent was asked to complete a questionnaire designed to provide additional information concerning viewpoints related to sonar contact reporting. This questionnaire was completed at the conclusion of the ranking tasks. The questions are reproduced in this section, together with response distributions and interpretive comments.

QUESTIONNAIRE

FOR EACH QUESTION, CHECK THE ONE RESPONSE THAT MOST ACCURATELY REFLECTS YOUR OWN OPINION.

1. During *present-day non-ASW peacetime operations* (for example, peacetime transiting, non-ASW refresher training, and so on), which do you feel sonar operators *should* do?

Officer	Enlisted		
8%	17%	a.	Report fewer false contacts than they do now even though it means increased delay in reporting.
38%	52%	b.	Continue present reporting practices.
54%	31%	c.	Report contacts more quickly than they do now even though it means more false contacts.

Question 1 is the only point in the survey at which officers and enlisted personnel differed significantly $[P(\chi^2) < .005]$. While the majority of officers favored "quicker" reporting during routine non-ASW operations, the majority of operators favored "continuing present reporting practices."

2. During *present-day ASW exercise situations*, which do you feel sonar operators *should* do?

    6% a. Report fewer false contacts than they do now even though it means increased delay in reporting.

    37% b. Continue present reporting practices.

    57% c. Report contacts more quickly than they do now even though it means more false contacts.

In Question 2, officer and enlisted personnel are in agreement: the majority favors "quicker" reporting. Since the response distribution is not significantly different from the officers' responses to Question 1, it is apparent that the enlisted personnel (as a group) changed their position in responding to the second question. The responses to these two questions may indicate some lack of communication from the officers to the sonar operators, since it is totally within the realm of these two groups to effect "quicker" reporting if that is desired.

3. Which have you found to be true?

LANT	PAC		
36%	55%	a.	Many more false contacts are reported during ASW exercises than during non-ASW peacetime operations.
32%	23%	b.	A few more false contacts are reported during ASW exercises than during non-ASW peacetime operations.
8%	2%	c.	About the same number of false contacts are reported during ASW exercises and non-ASW peacetime operations.
24%	20%	d.	Fewer false contacts are reported during ASW exercises than during non-ASW peacetime operations.

Question 3 is the only point in the survey at which the Atlantic Fleet and Pacific Fleet personnel differed significantly [$P(\chi^2) < .05$], with the latter indicating the ratio of ASW exercise to non-exercise false contact rate somewhat greater than the Atlantic Fleet experience. A possible explanation for this difference may be differences in the "non-ASW peacetime" environments of the fleets: the Atlantic Fleet has a higher rate of contact with "real" targets than the Pacific Fleet, reducing somewhat the contrast between "exercise" and "non-exercise" conditions, relative to the Pacific Fleet.

Note that very few felt that the *same* false contact rate obtained during exercise and non-exercise conditions. The 20-24% who responded "fewer false contacts during exercises" were probably thinking of situations in which knowledge of position of the target, or actual contact with it, decreased attention to non-target contacts. The large majority who responded "more" or "many more" false contacts during exercises are no doubt reflecting the effect of increased expectancy during exercises, resulting in increased false alarm rates.

4. How much of a problem do you think false contacts would be in the event of a war involving ASW?

18% a. Very serious problem.
33% b. Serious problem.
43% c. Moderate problem.
6% d. Insignificant problem.

In Question 4, the majority felt that false contacts would constitute a "serious" or "very serious" problem; very few felt that it would be "insignificant." No difference in response pattern existed between groups.

5. At the present time, how well do you think the bridge can judge the degree of certainty which sonar has about an initial contact?

3% a. Very accurately.
29% b. Accurately.
58% c. Not very accurately.
10% d. Not at all.

6. How useful do you think knowledge of sonar's confidence or certainty concerning an initial contact would be in decision making on the bridge during a war involving ASW?

83% a. Very useful.

14% b. Moderately useful.

2% c. Of little use.

1% d. Of no use.

Questions 5 and 6 indicate a potentially serious inefficiency in the communication of useful information from sonar to command. The majority of both officers and sonar operators (with no significant difference) felt that sonar's "certainty" judgments were "not very accurately" perceived on the bridge; yet the very large majority of officers and operators felt such information would be "very useful in decision making. The usefulness of such information in distinguishing false from valid contacts was substantiated by the detection experiment to be described.

7. From *your own experience* during present-day ASW exercises, what do you feel is a typical false contact rate

a. in coastal/shallow water?

________________ per ________________

number of false contacts *unit of time*

b. in mid-ocean/deep water?

________________ per ________________

number of false contacts *unit of time*

8. *In the event of war* involving ASW, what do you feel a typical false contact rate might be

a. in coastal/shallow water?

________________ per ________________

number of false contacts *unit of time*

b. in mid-ocean/deep water?

________________ ________________

number of false contacts *unit of time*

The responses to Questions 7 and 8, whose formats are shown above, are depicted for the total sample (there were no differences between groups) in Figure 13 in the form of

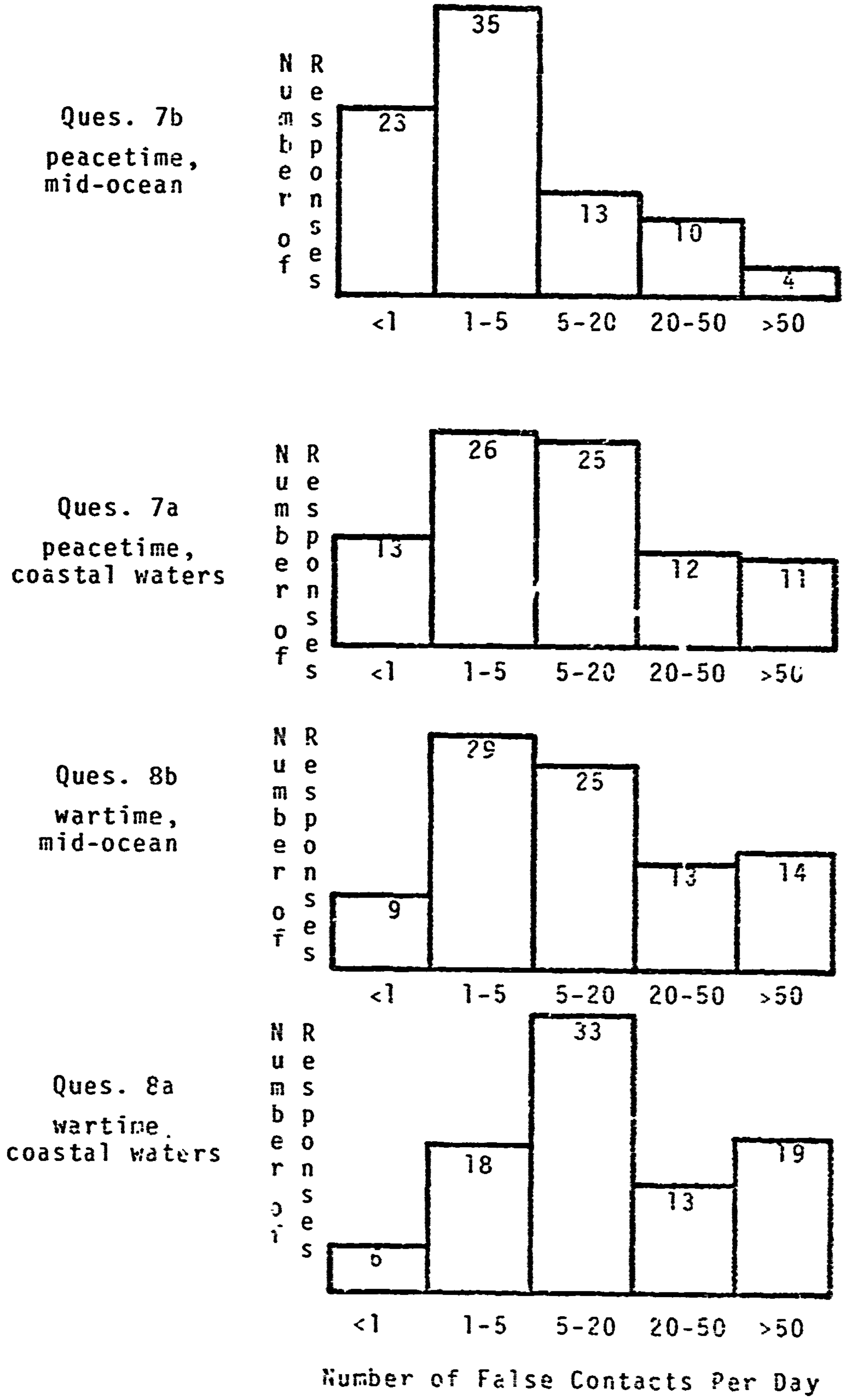


Figure 13. Distributions of false contact rate estimates.

histograms. The histograms have been arranged in order of increasing estimates of false alarm rate from top to bottom; the difference between the distribution of responses for Questions 7b and 8b is very significant [$P(\chi^2) << .001$]. The difference between mid-ocean and coastal estimates is partly due to differences in sonar environment; the difference between exercise and wartime estimates no doubt reflects the expected effects of wartime decision consequences in lowering the detection/reporting threshold.

Questions 7 and 8 were intended more to measure the opinions of the personnel sampled than to project actual false alarm rates; but if these opinions proved to be accurate, it is obvious that wartime false contact rate would be very substantial indeed, posing a very significant attack decision problem.

CHAPTER III

THE DETECTION EXPERIMENT

Objective

The objective of the detection experiment was to compare sonar operator detection and reporting performance under two conditions that would be perceived by the operators as having very different consequences as a direct function of their performance. The problem is to identify controllable variables associated with such conditions. Ideally, one would like to manipulate perceived threat directly. Even with ethical questions aside, it is doubtful whether this can ever be done while maintaining sufficient experimental control to make the results meaningful. There is one operationally meaningful variable, however, that relates directly to perceived importance of the tactical situation if not to perceived threat *per se*. That variable is *"command attention."*

Because of their relative isolation from command and control activities, often for very long periods of time during non-ASW evolutions, sonar operators appear to be particularly responsive to communications from command that bear upon the importance of their performance. Such communications may occur with very different frequencies as a function of the tactical situation and individual command attitudes. However, it is widely assumed in the Navy that high levels of proficiency among ASW ships are typically associated with high levels of "command attention."

It seems reasonable to assume that, in general, "command attention" will vary with the tactical situation; it will be very low during non-ASW missions, moderately high during ASW competitive exercises, and highest of all in the event of a true threat. It was our objective to set up two very different situations with respect to the evident "command attention" and to measure the resulting differences in detection and reporting performance by typical sonar operators.

We felt that the effect of "command attention" on performance could manifest itself in the setting of the *reporting criterion*, or threshold; but we also felt that it could affect *detection efficiency*, or effective signal detectability. The distinction between these two aspects of detection behavior is one of the important contributions of signal detection theory to psychology.

We intended to measure operator detection performance against a given set of signals in a setting which would first minimize apparent command attention; and measure performance of the *same* operators against the *same* stimuli in a setting in which maximum possible command attention would be brought to bear. The intent was not to explicitly manipulate the reporting criterion, but to create two situations *genuinely* perceived by the operators to be of very different importance with respect to their detection performance. We wished to use the AN/SQS-26 sonar as the vehicle for this experiment because of its status as the Navy's principal long-range active sonar.

Method

To accomplish the objectives set forth, it was necessary to place a number of fleet sonar operators in a credible setting of high "command attention" while providing a realistic sonar display. This requirement ruled out an at-sea experiment, since realistic detection situations for several operators could not be set up without prohibitive requirements for ship services; and, in any event, experimental control of many important variables would be completely lacking, leading to "unrepeatability" which would prevent meaningful comparative analyses of the data. Performing the experiment employing a shipboard sonar at the dock was not feasible either, because the only method of realistically activating the sonar would have been with tape recordings of signals and reverberations, which were not then available for the AN/SQS-26 sonar. The signal injection equipment which is part of the sonar set could not be used because that equipment is mounted in the sonar control room (with the display consoles), so that no covert way was available to manipulate the controls necessary to simulate targets.

The nature of the AN/SQS-26 A-scan display, however, permitted a very satisfactory experimental approach. This display, unlike older sonar displays, is a memory-type device which presents a history of up to the "last six pings" in a *static* presentation. This static characteristic permitted the use of still photography to create a realistic representation of the sonar display. We photographed an AN/SQS-26 A-scan display after *each* ping was "painted" onto the display, then projected the resulting ping-by-ping sequences of color transparencies by rear projection techniques onto translucent display screens of exactly the same size as the sonar display CRT. The resulting display was very similar to the presentation observed at the sonar; in fact, none of the operators tested felt it was different in any significant way from the presentations on their own equipment.

Details concerning the stimulus materials, the apparatus, the subjects, the procedure, and the experimental results are given in the following sections.

Stimulus Materials

An AN/SQS-26 CX located at the Fleet ASW School, San Diego, was used to prepare the stimulus materials. The target simulation capability incorporated in the sonar (Unit 34), which provides selection of simulated target course, speed, range, bearing, echo intensity, and own-ship heading and speed, was suitable for generating targets for the detection task. However, the noise generator which is a part of this unit proved unsatisfactory as a source for noise/reverberation simulation, since the same voltage waveform is introduced into all 24 sonar channels, providing total spatial coherency. This produces very obvious noise "bands" across the entire sonar display, and is easily detected as being artificial. Therefore, we constructed a special AN/SQS-26 noise/reverberation generator. The device was designed to provide broadband Gaussian noise across the sonar bandwidth, and, for reverberation simulation, white Gaussian noise passed through a relatively broad bandpass filter for FM (also called coded pulse, or CP) channel reverberation, and a very narrow bandpass filter for CW channel reverberation. A solid state active filter system was designed and a prototype was constructed and tested for a single beam of the sonar. The prototype design proved to be very satisfactory, so the final unit was constructed. This unit as 12 independent solid state noise generators, 12 active bandpass filters for CP reverberation simulation, 12 active bandpass filters for CW reverberation simulation, and operational amplifiers for the linear combination of the various noise/reverberation outputs. While each sonar beam receives its primary energy from the associated channel within the noise generator, it also receives some energy from adjacent channels of the generator, simulating the limited spatial coherency evident in actual sonar operation, which results from overlapping beams; the necessary mixing is done within the noise/reverberation generator. The 12 outputs of the generator, each containing the CP and CW backgrounds for a given beam, are fed into the 12 post-amplifiers of the sonar set which immediately follow vertical and horizontal beam forming. A schematic diagram of the noise/reverberation generator is given in Appendix C.

With the noise/reverberation generator connected to the sonar, the A-scan display was adjusted according to NUSC/New

London Technical Memorandum 2134-540-70. All transparencies of the A-scan display were photographed with a 20-Kyd zone width and with the sonar in the single storage mode, which permitted a ping-by-ping history to build up from zero to six pings. It also permitted use of one storage tube only, and optimization of display adjustments. The photographs were taken with a Nikon-F camera and Micro-NIKKOR Auto F3.5, 55mm lens with settings of f5.6 and 1/8 second on Kodak High Speed Ektachrome film from a distance of approximately 30 inches, and were normally processed. A four-digit light-emitting-diode numerical display was mounted on the lower right-hand portion of the sonar CRT, which permitted recording the sequence number (1 to 72) and the ping number (1 to 6) in the image on each slide. Sequences of 6 consecutive pings were photographed in the following way: The display was erased, and the sequence/ping number counter was initialized. The test target, if one was to be present, was inserted with the proper parameters for the first ping, and the sonar was allowed to cycle normally. After the first ping was fully recorded on the display, the first photograph was taken. During the sonar dwell-time, the target echo intensity was adjusted to the proper value for the second ping, according to a schedule of random signal intensities. After the second ping had fully developed, the second photograph was taken; and so on. Thus, each photograph represented the sonar display as it actually appeared after a given number of pings. This process was continued until the sixth ping was completed and photographed. A sequence of six slides thus depicted the sequential development of six simulated sonar pings.

A total of 72 six-ping sequences was photographed, 24 sequences with noise/reverberation background only; 24 with the noise/reverberation background plus an FM (or CP) target; and 24 with the noise/reverberation background plus a CW target.

For the sequences containing targets, target ranges and bearings were randomized to approximate a uniform distribution over the sonar display; and target range-rate was randomized to approximate a uniform distribution within the following intervals: 25 to 15 knots closing; 5 knots closing to 5 knots opening; and 15 to 25 knots opening. Within each signal sequence, the signal level was randomized from ping to ping with an approximate normal distribution with 3.5 db standard deviation, which is representative of actual AN/SQS-26 operation; and with a mean signal-to-noise ratio approximately equal to the minimum detectable level (MDL) for the particular signal type (CP or CW). Thus, many pings within a signal sequence typically did not contain *perceptible* echoes, even though a signal had been injected. These parameters realistically simulated "difficult-to-detect" targets.

After developing, the resulting slides were inserted into Kodak Carousel slide trays, 9 sequences (54 slides) to a tray, for a total of 8 slide trays.

Apparatus

Five rear-projection display screens were constructed with 10-3/4" wide by 11" high viewing areas (approximately the same size as the CRT sonar display) mounted in 16" wide by 18" high frames with range scales on the right and left sides and "beam number" scales at the top and bottom to permit operators to estimate the range and bearing of a target.

Five Kodak 850 Autofocus slide projectors with zoom lenses were employed to project the sonar display images. The automatic focus feature kept the slides in focus automatically, and the zoom feature permitted adjustment of the images to the same scale as the original sonar display. The five projectors were connected to a pair of Hunter Model 111-8 timers, which were connected in flip-flop configuration to provide 300 millisecond advance pulses (actually, contact closures) simultaneously to the five projectors every 25 seconds.

Response buttons were provided each operator for reporting sonar contacts. Each response button was connected by wire via an "interface box" to a single tape recorder, which was run ring each experimental session. The operators were told that e ping number corresponding to their push button response was being recorded electronically, to discourage claiming "early" detections by writing in ping numbers on their pen-and-paper response sheets that occurred earlier than the point in time at which the reporting decision was actually made. In fact, data on response time were not actually recorded electronically.

Subjects

Twenty U. S. Navy sonar technicians from AN/SQS-26 ships at Newport, Rhode Island, served as subjects. Two of these were transferred before completing the experiment, and two more were found to have no operational experience with the AN/SQS-26 A-scan display. The remaining 16 were qualified watchstanders or watch supervisors. Six had 1 year or less of shipboard experience, while 10 had more than a year of experience. The 16 men included the following rates. four STG-2; ten STG-3; two STG/SN. As a group, their qualifications appeared typical of present-day shipboard sonar personnel.

Procedure

The detection experiment was arranged to be conducted in Newport, Rhode Island, through the cooperation of Commander, Destroyer Development Group. Excellent experimental facilities were provided us in a room at the Naval Underwater Systems Center, Newport. The windows were blacked out to permit low-level artificial illumination during the experiment, and five experiment booths were placed in the room in such a manner that each subject sat in a cubicle which prevented his seeing other sonarmen or their displays. Each sonarman sat in a student's chair, which provided him with a writing surface, and the rear-projection screens were placed on small tables immediately in front of him. The slide projectors were set up side-by-side across the room from the five adjacent subject booths, and the projector remote-control cables were all connected to the timing apparatus situated at the experimenter's station.

Five sonarmen performed the detection task at a time. All the operators were given their first session during a single week, and their second session during the following week. Arrangements for sonarmen to serve as subjects, and for their transportation to and from the experimental site, were provided in a very effective manner by Destroyer Development Group.

Prior to commencing the detection task during the first week, the instructions reproduced in Appendix D were read to the sonarmen, and sample photographic transparencies were shown to demonstrate the display and the sonar contact reporting procedure. This procedure (which may be best understood by reading Appendix D) called for a sonar contact report only when the sonarman felt he recognized a contact. This method was felt to most realistically approximate the actual sonar detection/reporting task (in contrast to the frequently imposed forced-choice situation associated with detection theory experiments, which requires responses at fixed intervals). The sonarmen were instructed that when they wished to report a contact, they should immediately press the response push botton, then record certain pertinent data concerning the contact on one of the response sheets (reproduced in Appendix E). They were to record the sequence number and ping number on which they responded, whether the target was on the CP or CW portions of the display (or both), the estimated range of the contact, the beam number, and whether the target appeared to have an opening, closing, or zero range rate. More than one report per sequence could be made; or none could be made. The operator was also required to evaluate the quality of each contact he reported *at the end of the sequence containing that contact.* This was done by indicating whether the contact was judged to be "good," "fair," or "marginal." He was also permitted to revise his contact judgments after all six pings had been presented; this on occasion led to the conclusion that an early report was in fact "not a contact."

The slides were advanced at 25-second intervals (slightly faster than real time) and were presented continuously within blocks of 18 six-ping sequences, taking 45 minutes; there were five-minute breaks between blocks, during which the subjects could move about. The total duration of the detection task was about 3 hours plus time for instruction of the subjects and for breaks. All subjects completed the task during the morning hours (0800 to 1130) of normal working days.

As can be seen from Appendix D, during the first session (Condition I), the operators were told that the purpose of the task they were to perform was to learn "what constitutes a reportable contact with the AN/SQS-26 sonar." No uniformed Navy officers or enlisted personnel were in the vicinity, and only the civilian experimenter and the subjects were present within the experiment room. Thus, while the environment for each sonarman's first session was certainly novel for him in some respects, we think these novelties can be characterized as those of an "experiment," conducted by civilian researchers, having no clear connection with his accustomed environment (the operational Navy); and without clear consequences for either errors or error-free performance of the task. The origin of the stimulus materials was never discussed with the sonarmen, and knowledge of results was never given.

In contrast, when the sonarmen arrived for their second session during the subsequent week (Condition II), a U. S.

Navy Commander, in uniform, was present in the room. When the sonarmen were all present and their identities verified, they were addressed by this officer. He made a statement to the men (from memory) following this guideline:

> The exercise last week was, in reality, just a warm-up for the very important task I want you to perform today. In the interest of developing optimum tactics against the Soviet submarine threat, you are being asked to view a series of recently acquired data where it is known that a number of targets may be detectable. Security prevents my discussing the nature of these data in detail, but it is extremely important that every valid contact be identified as such. Our estimate of the threat clearly depends on the most accurate information we can get in this respect. Therefore, please report any contact that you feel qualifies as a "possible sub." Report at the earliest point in the sequence where you feel such a report should be made.

Following this statement, the men were seated in their booths and given "Condition II" response booklets. (The Commander was not present during the actual detection task.) These were of the same format as the "Condition I" booklets with three very apparent exceptions: the pages were pink rather than white; the pages were conspicuously marked "SECRET NOFORN when any identifier block completed"; and each page had this "identification block" (which was never in any case completed):

IDENTIFIER BLOCK Do not write in this space	Contact Code_________(1)	p(D)___________(4)
	Threat Code__________(2)	p(FA)___________(5)
	Type_________________(3)	Z-Code__________(6)

A sample page from "Condition II" response booklet is shown in Appendix F.

Therefore, the setting for the detection task in the second session was quite different from the first. Every effort was made to make performance of the detection task appear relevant and consequential to the "operational Navy" via the "command attention" focused upon it. (Hopefully, this was not a deception--we certainly intended this research to be "relevant and consequential!") The intent was to create a detection situation which, while not in an at-sea tactical environment, was *genuinely* perceived by the operators to be analogous, and not "merely" a research exercise. We believe that the photographic reproductions of the display, which could have been recorded during a tactically significant encounter (an explanation for their origin which was at least implicit in the second session), and the delivery of the Condition II instructions by a Navy Commander, did indeed set the scene as we desired.

Independent Variables

We review and summarize the independent experimental variables here, and the criterion variables in the next section, as a prelude to description of the experiment results.

Command Attention. Two levels, "low" and "high," were administered during each subject's first and second experimental sessions, respectively, as previously described. It should be noted that this experimental design is not "balanced" with respect to the "command attention" variable, in that the two levels of this variable were always administered in the same sequence. In general, this design invites the confounding of such sequence effects as "learning" with the principal experimental variable. The nature of the "command attention" variable, however, does not at all lend itself to administration in the opposite sequence; that is, administration of the "high command attention" level during the first session would make it doubtful that a satisfactory control condition could be achieved during the second session. For this particular experiment, however, we felt that "learning" effects would have a negligible influence, for at least two reasons: (1)

each subject was an experienced AN/SQS-26 operator who was quite familiar with the information display employed in the experiment; and (2) no feedback of performance results was ever provided.

Signal Type. Equal numbers of FM (also called CP) and CW simulated targets were represented in the stimuli, which appear on separate parts of the CRT sonar display. Since the character of these two types of signals, as displayed, is very different (principally owing to processing bandwidth differences), it is reasonable to expect differences in performance at detecting these signals. Therefore, the sonar contact reporting procedure was designed to permit differentiation of "FM" and "CW" targets.

Time Blocks. Each 3-hour experimental session was divided into four 45-minute "time blocks" by 5-minute breaks, as previously described. The experimental data were analyzed with respect to time blocks to reveal possible time effects on performance.

Criterion Variables

Correct and False Reports. Each sonar contact report was scored "correct" if [illegible] met the following criteria:

1. The ping sequence during which the report was made must have been one of those 48 in which a simulated target was injected.
2. The report must have indicated the correct signal type (FM or CW).
3. The estimated range of the target must have been within 5 Kyds of the "true" range of the target (i.e., the range at which it was injected) within the displayed zone.
4. The sonar beam which the target was estimated to be in must have been within 3 beams of its "true" beam (i.e., where it was originally injected).

If a sonar contact report failed to meet any of these criteria, it was scored as a "false" report.

It was expected that range and bearing estimation precision would be considerably better than the allowed tolerances (which it proved to be), even without the aid of the sonar range/bearing cursor, which was not available to the operators. These tolerances were selected to provide a generous interval for the range/beam accuracy analysis discussed below; the four criteria, together, were sufficiently strict to provide the necessary discrimination of "correct" from "false" reports.

Reported Quality Level. It was expected to be of interest to compare the quality judgments for "correct" target detection reports with those for "false" reports, as an indicator of the diagnostic value of the judgments in distinguishing valid targets.

Range and Beam Estimation Accuracy. For each correct report, the differences between the estimated range and beam and the "true" range and beam of the injected target were calculated to be used as additional performance measures.

Ping Number. Because latency of sonar contact reporting is an operationally important variable (see "Results: Consequences of Delayed Contact Reporting" in the previous chapter), all sonar contact reports made during the experiment were identified by the ping number on which they occurred. (All six-ping sequences which contained targets had target-plus-background present for all six pings, but because a *particular* sequence represented a realization of a normal random process--with mean signal-to-noise ratio at the minimum detectable level, and a standard deviation of 3.5 db--the signal may in fact have been undetectable for several pings· therefore, reporting latency must be considered across *all* sequences to achieve a statistically meaningful latency measure which combines operator, signal, and noise random processes.)

Results

In this section, the effects of the independent experimental variables are analyzed and described in relation to each of the criterion variables in turn.

Effects on Number of Correct and False Reports. A three-factor analysis of variance was calculated for the number of correct reports and the number of false reports, based upon the 2 (levels of command attention) x 2 (FM or CW signal type) x 4 (time blocks) within-subjects experimental design. A summary of the analysis of variance for correct reports is shown in Table 3, and for false reports in Table 4. It can be seen that for *correct* reports, "signal type" had a significant main effect ($p < .05$) and the "command attention" x "signal type" interaction had a significant effect ($p < .05$). For *false* reports, it can be seen that "command attention" had a significant main effect ($p < .01$) and "signal type" had a significant main effect ($p < .01$). "Time Blocks" had no statistically significant effects in this experiment.

Figures 14 and 15 show the total numbers of *correct* and *false* reports obtained in the experiment, respectively, classified by the two independent variables revealed by the analyses of variance to be significant: "command attention" and "signal type." Figure 14 shows that there were more CW correct reports than FM correct reports (recall that equal numbers of detection opportunities were presented for each signal type), and that the "high" level of command attention resulted in a very slight decrease in the number of correct FM reports, but a *substantial increase* in the number of correct CW reports, compared to the "low" level of command attention. This differential effect of "command attention" on FM and CW correct reports, of course, resulted in the significant "CA x ST" interaction shown in Table 3, but caused the "CA" main effect to fall short of the statistical significance criterion.

TABLE 3

SUMMARY OF ANALYSIS OF VARIANCE FOR CORRECT SONAR CONTACT REPORTS

SOURCE	df	MS	F	p
S (Subjects)	17	14.05		
CA (Cmnd Attn)	1	2.92	1.22	
CA x S	17	2.39		
ST (Signal Type)	1	45.92	6.79	p<.05
ST x S	17	6.76		
TB (Time Blocks)	3	0.90	0.39	
TB x S	51	2.30		
CA x ST	1	5.28	5.56	p<.05
CA x ST x S	17	0.95		
CA x TB	3	0.35	0.37	
CA x TB x S	51	0.94		
ST x TB	3	4.09	2.21	
ST x TB x S	51	1.85		
CA x ST x TB	3	0.62	1.32	
CA x ST x TB x S	51	0.47		
Total	287			

TABLE 4

SUMMARY OF ANALYSIS OF VARIANCE FOR FALSE SONAR CONTACT REPORTS

SOURCE	df	MS	F	p
S (Subjects)	17	61.79		
CA (Cmnd Attn)	1	66.12	11.96	p<.01
CA x S	17	5.53		
ST (Signal Type)	1	425.34	12.43	p<.01
ST x S	17	34.22		
TB (Time Blocks)	3	0.06	0.02	
TB x S	51	3.10		
CA x ST	1	7.34	1.51	
CA x ST x S	17	4.86		
CA x TB	3	0.46	0.47	
CA x TB x S	51	0.98		
ST x TB	3	3.91	1.13	
ST x TB x S	51	3.45		
CA x ST x TB	3	1.80	1.55	
CA x ST x TB x S	51	1.16		
Total	287			

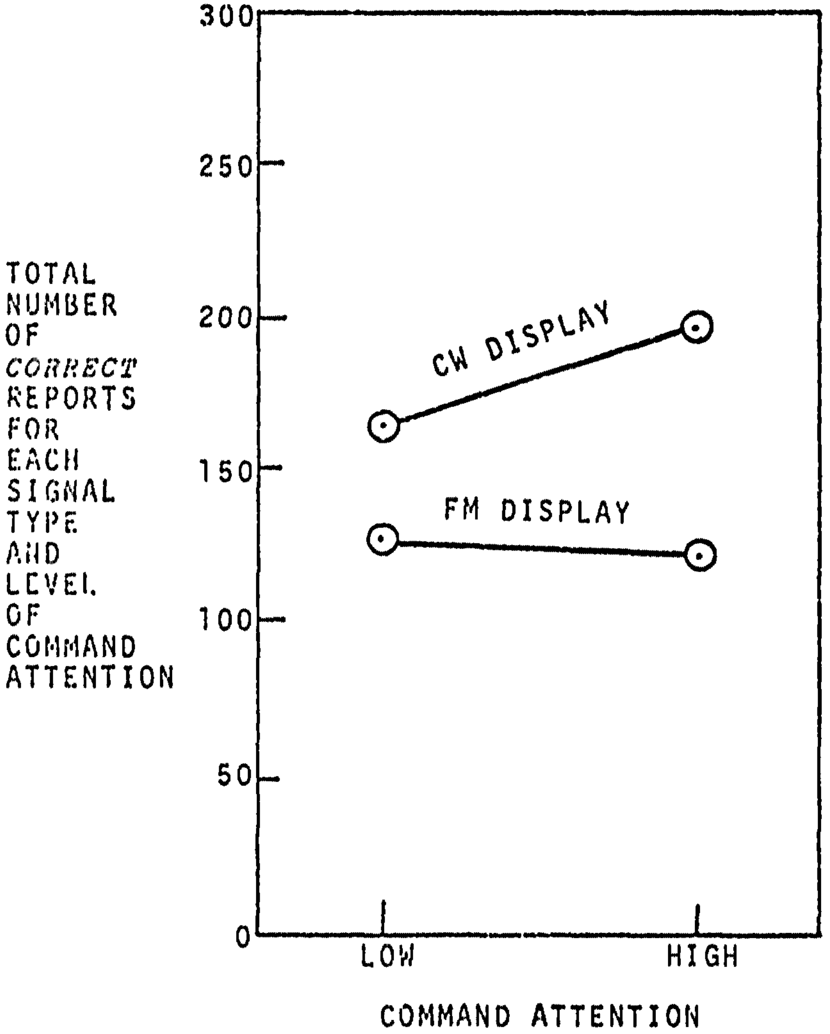


Figure 14. Total number of *correct* reports by signal type and level of command attention.

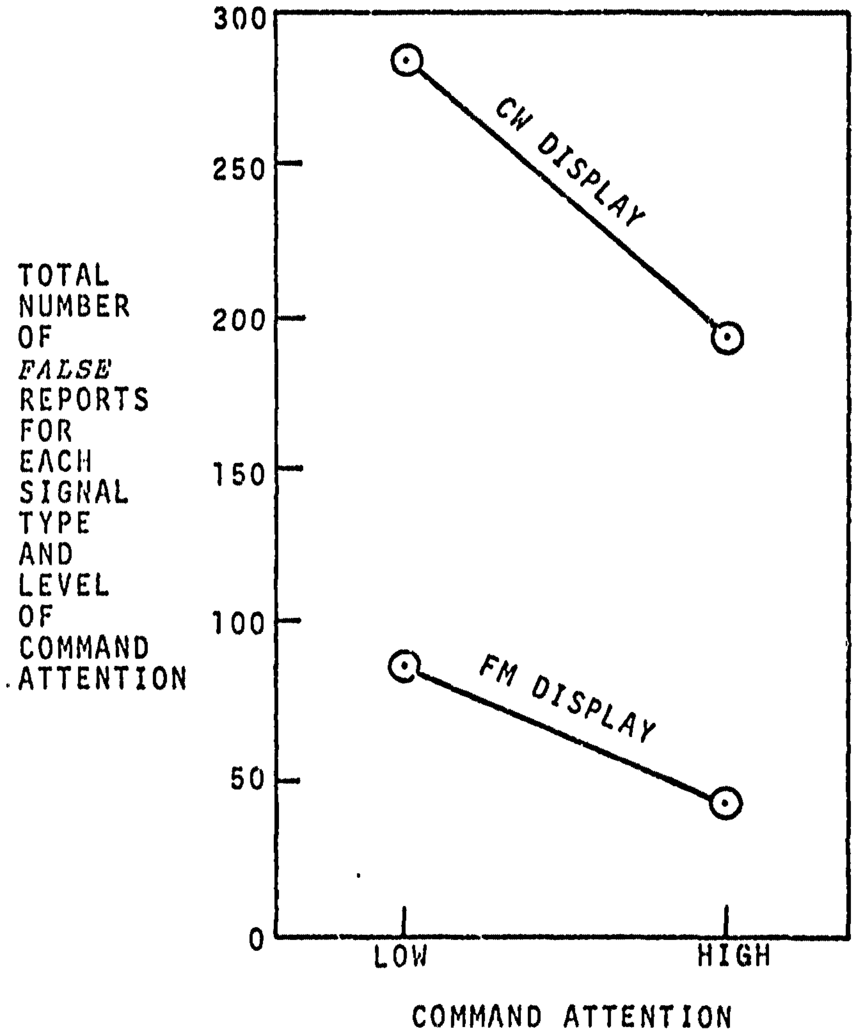


Figure 15. Total number of *false* reports by signal type and level of command attention.

Figure 15 shows that there were many more CW false reports than FM false reports, and that the "high" level of command attention *substantially reduced* the numbers of both CW and FM false reports, compared to the "low" level of command attention. Since the "high" level of command attention reduced *both* CW and FM false reports, the "CA x ST" interaction was not significant for false reports (Table 4).

Thus, the "high" level of command attention resulted in a *significant improvement in detection performance*, compared to the "low" level, by reducing FM false reports by 50% while reducing FM correct reports by only 4%; and by reducing CW false reports by 31% while *increasing* CW correct reports by 20%. The differential effect by signal type is attributed to the different psychophysical tasks presented by the FM and CW portions of the display.

Effects on the Diagnostic Value of "Contact Quality" Judgments. Figure 16 shows the distribution of correct and false reports among the "contact quality" categories. (Recall that a "quality" judgment was required for each contact report the operators made.) It can be seen from the distribution of all *correct* reports among the "quality" categories that the most likely quality judgment for reports elicited by genuine targets was the highest level, "good," containing approximately half of all correct reports made. The proportions of correct reports given lesser quality judgments decreased monotonically with the decreasing "quality" descriptors.

On the other hand, it can be seen from the distribution of all *false* reports among the "quality" categories that the most likely quality judgment for reports elicited by the noise/reverberation background was the second highest level, "fair," containing approximately 40% of all false reports made.

Thus, the distributions shown in Figure 16 indicate that the subjects' "contact quality" judgments had diagnostic value

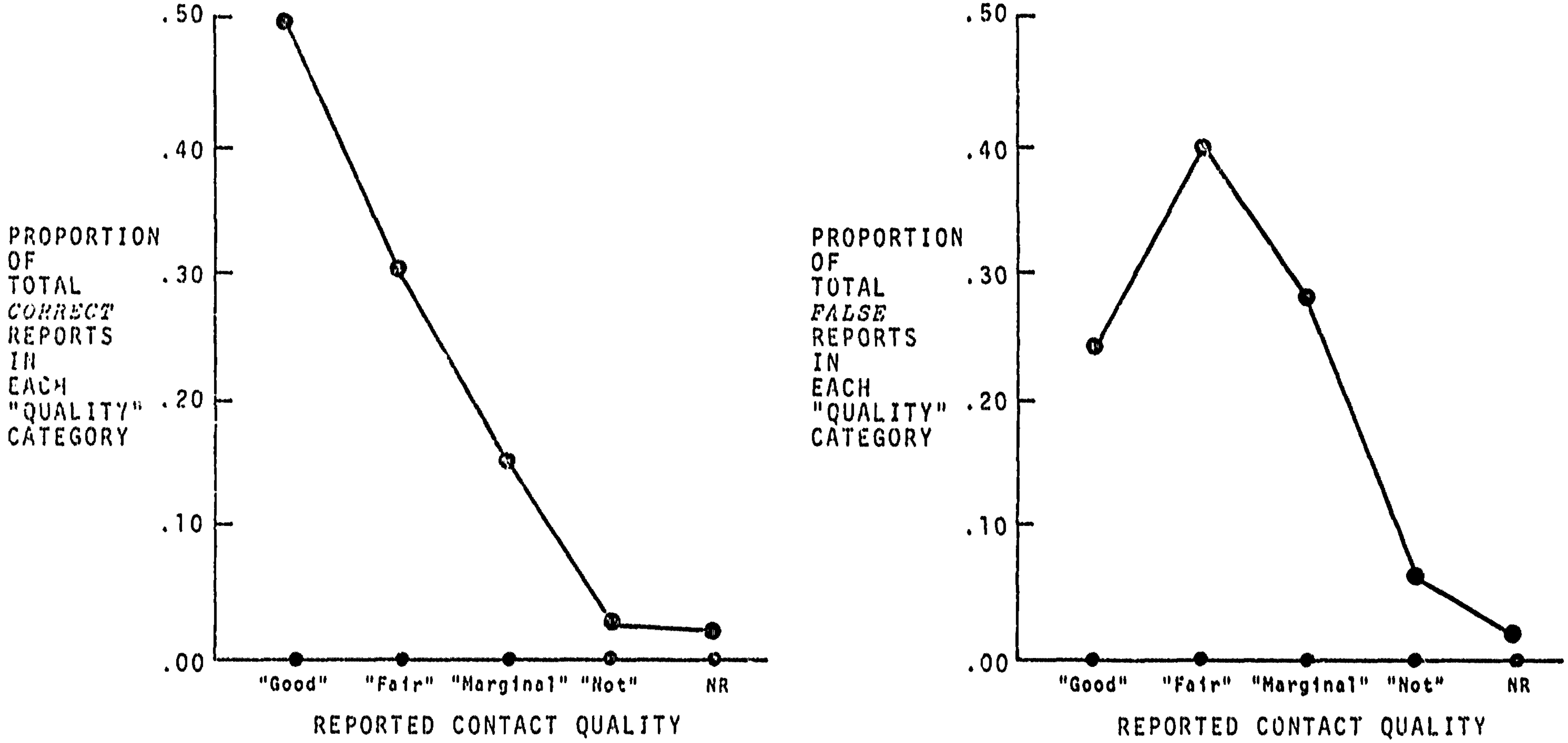


Figure 16. Distribution of correct and false reports among the "contact quality" categories.

in discriminating correct from false reports. This diagnostic potential may be examined more directly by considering the proportion of all (i.e., correct plus false) reports in a given quality category which proved to be correct. For example, for the "low" level of command attention, there were 62 FM targets reported at the highest quality level ("good"), 46 of which were scored correct. Therefore, the proportion of correct reports for the specified conditions was 46/62 = 0.74. Figure 17 shows the proportions of reports in each quality category which proved to be correct, by signal type, reported contact quality, and level of command attention.

It can be seen from Figure 17 that, without exception, the proportion of correct reports increased monotonically with increasing perceived quality levels. For example, for FM reports made during the "low" command attention condition, 48% of the contacts judged to be of "marginal" quality proved to be correct, 58% of the "fair" reports proved to be correct, and 74% of the "good" reports proved to be correct.

The proportion of correct FM reports in each quality category is seen to be greater in every case than the corresponding proportion of correct CW reports. This is the result of more "strict" reporting criteria employed by the operators for the FM display, an explanation substantiated by the fact that, while equal numbers of FM and CW targets were presented, fewer correct (and false) FM reports than CW reports were made during the experiment.

It can also be seen from Figure 17 that the "high" level of command attention resulted in improved target discrimination performance in every contact quality category, for both types of signals. For example, for FM targets in the "good" category, the proportion of correct reports was increased from 74% for the "low" level of command attention to 84% for the "high" level. Similar improvements are seen in the other

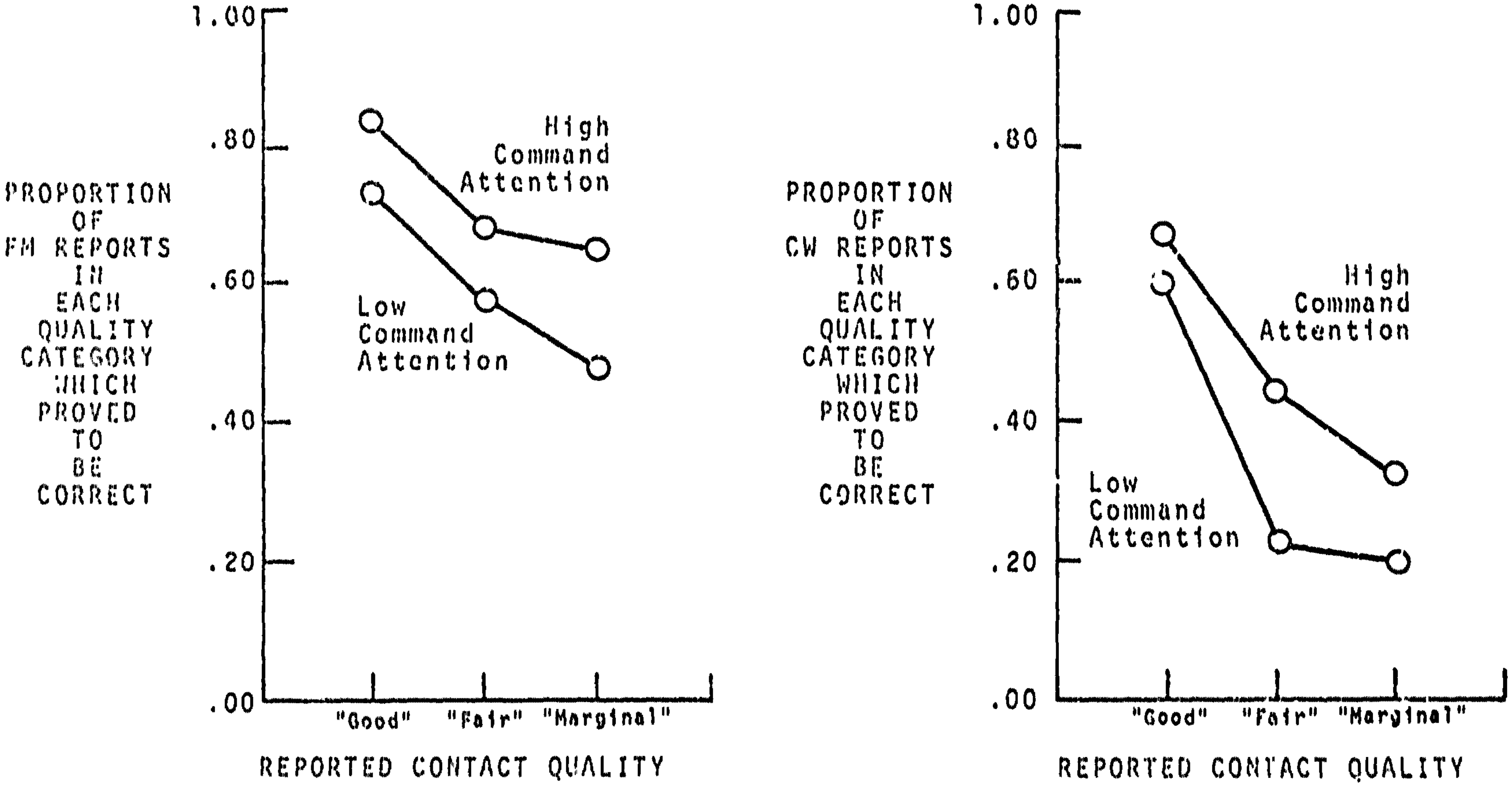


Figure 17. Proportion of reports in each quality category which proved to be correct, by signal type, reported contact quality, and level of command attention.

categories. This improvement in target discrimination performance resulted from the employment of more "strict" reporting criteria during the "high" command attention condition. Generally, (when "detection efficiency," or effective signal detectability, is constant), *the penalty paid for an improvement in discrimination performance (which must be obtained through a shift to a more "strict" or "cautious" reporting criterion) is a decreased probability of detection*, resulting from the consequent decrease in numbers of false *and correct* reports. In this case, however, the increase in target discrimination performance brought about by the "high" level of command attention was accompanied by only a 4% decrease in the number of correct FM reports, *and by a 20% increase in correct CW reports*, compared to the "low" command attention condition. Thus, it must be concluded that detection efficiency, or effective signal detectability for the display-operator system, was greater for the "high" level of command attention than for the "low" level. This, of course, is a very desirable result; it is considered in greater detail in the section on "Signal Detection Theory Analysis."

Effects on Range and Beam Estimation. Another indication of improved performance during the "high" command attention condition may be seen in Figure 18. This figure shows the means of the absolute values of errors in estimating the range to valid targets (i.e., $n^{-1}\Sigma|R_{true}-R_{estimated}|$) for the two levels of command attention. A significant decrease ($z=3.3$, $p < .001$) in mean range estimation error was obtained for the "high" level of command attention, compared to the "low" level, representing a 19% improvement in target range estimation performance. (Note that while these range estimation errors are useful measures of performance between the two command attention conditions of this experiment, their *actual values* are not to be taken as estimates of AN/SQS-26 range accuracy, since the sonarmen *did not* have the electronic range cursor normally used for this purpose.)

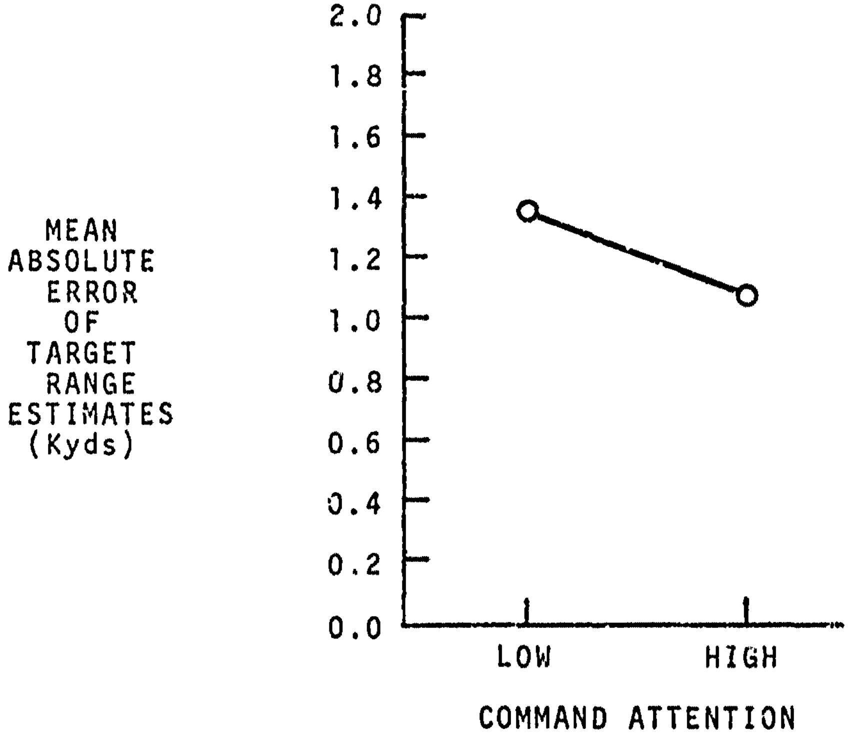


Figure 18. Mean absolute error of target range estimates by level of command attention.

While range estimation was performed on a continuous scale from zero to 20,000 yards, target *beam* estimation was a rather "coarse," discrete task offering only 12 response possibilities. Mean absolute beam estimation accuracy during the "low" command attention condition was 0.71 beam, which left little room for improvement. Mean absolute beam accuracy during the "high" command attention condition was 0.70 beam, not significantly different.

Distribution of Reports by Ping Number. Figure 19 shows the distribution of all sonar contact reports by the ping number (1-6) on which they were made. There were no meaningful differences in this distribution by level of command attention, signal type, or correct versus false reports. However, it should be noted that since each target sequence presented six "echoes" with uncorrelated, normally distributed signal intensity probabilities, a single target sequence can be regarded as a single realization of a six-variate uncorrelated normal probability distribution (with equal means μ=MDL and equal standard deviations σ=3.5 db). Therefore, the 48 target sequences do not represent a very large sample (of a six-variate joint probability distribution) and do not permit great precision in looking for differences in reporting distributions by ping number. Thus, one cannot confidently conclude from these data that there are none.

It can be seen from Figure 19 that, with one exception, the proportion of contact reports increased monotonically with the ping number. Very few reports were made on the first and second pings; the majority of reports (84%) occurred on pings 4, 5, and 6. This result was expected, since multiple observations are an extremely important factor in sonar detection. This fact is the principal reason for providing a memory-type display, and its utilization is reflected in a survey of SQS-26 operators, in which "echo consistency" (over multiple observations) was the clue most frequently indicated to be of

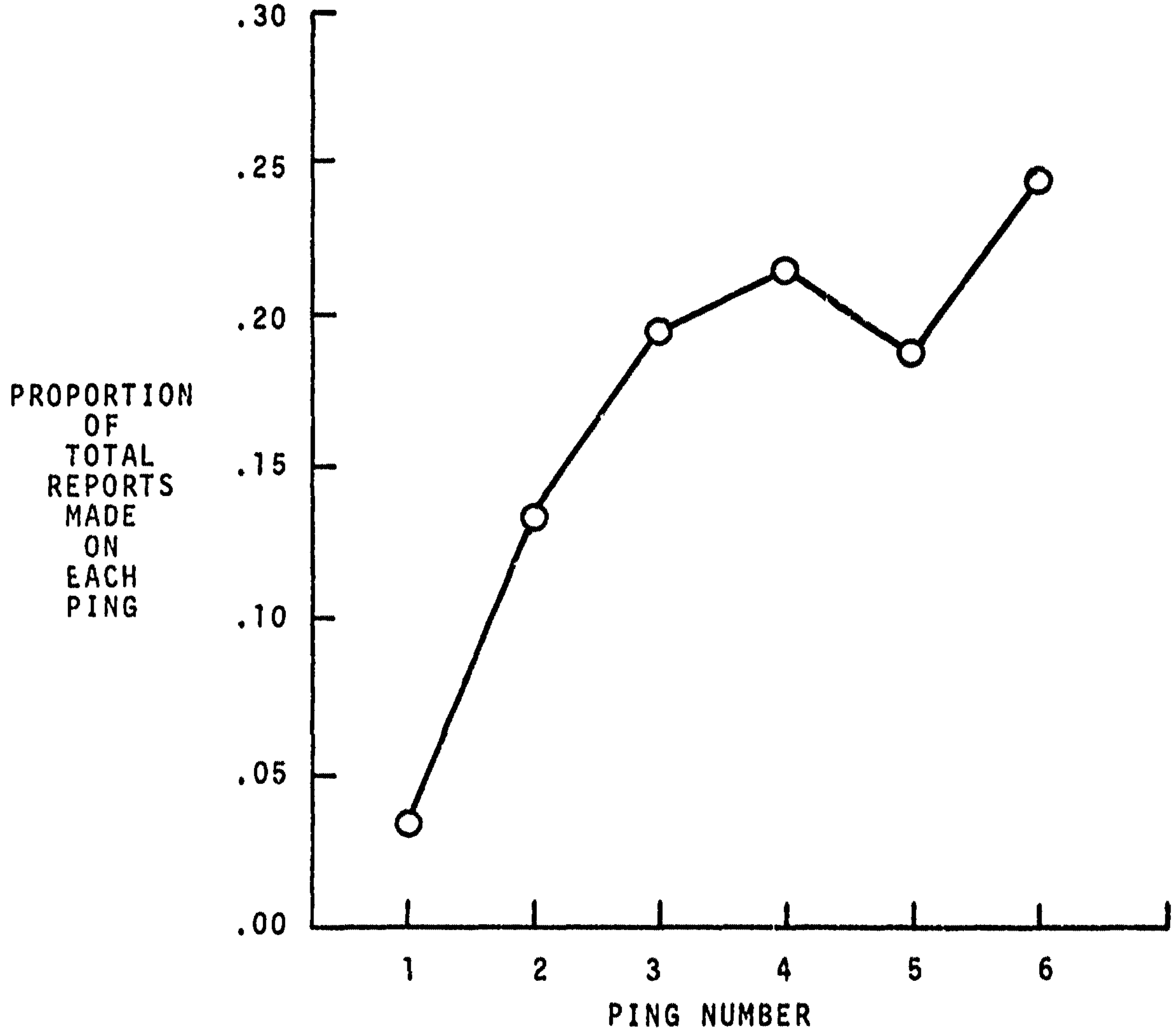


Figure 19. Distribution of sonar contact reports by ping number.

importance in target detection (Stern, 1971). We feel that the sonar detection task is best modeled as a *sequential analysis task* (see Wald, 1947; Birdsall and Roberts, 1971), in which additional observations may be made if the evidence at hand does not justify a detection report. This model prescribes comparatively strict reporting criteria for decisions based on few observations, becoming less strict as the number of observations increases, as a result of the greater statistical significance attached to likelihood estimates based upon greater numbers of observations. Experimental confirmation of the use of reporting criteria by sonar operators which vary with the number of observations has been obtained for passive sonar (Harabedian *et al.*, 1970) and for active sonar (Rizy, 1972). The data shown in Figure 19 substantiate these findings.

The exception to monotonically increasing proportions in Figure 19 occurs on ping 5, which we suggest is owing to its position as "next-to-last," and is thus an experimental artifact. On ping 5, the sonarmen had a slight bias toward taking that "last" look. If ping 6 were not "last," the proportion of reports made on ping 5 probably would not have been anomalous.

Signal Detection Theory Analysis

The results presented in the last section were based directly upon effects seen in the experiment criterion variables, without transformations of any kind. However, it is also desirable to describe the results in terms of probability of detection, p(D), probability of false report, p(FR), the receiver operating characteristic (ROC), and the signal detection theory statistics d′ and β. These descriptions do require transformations of the direct experimental data, which involve certain underlying assumptions. We will attempt to clearly identify each of these assumptions as they are made, so the reader may use the transformed data in a completely informed way.

Probability of Detection, p(D). An estimator of p(D) for the stimuli employed in this experiment may be formed from the ratio of the number of correct reports (for a given set of independent variables) to the number of detection opportunities (for that set of variables). There were 24 sequences in which an FM target was injected, and there were 18 operators, so there were a total of 432 detection opportunities for FM targets for each level of command attention. Likewise, there were 24 CW target sequences, so there were 432 CW target detection opportunities. Thus, we may form estimates for p(D) for the experiment from the observed total numbers of correct reports:

COMMAND ATTENTION	SIGNAL TYPE	NUMBER CORRECT REPORTS	OPPORTUNITIES	p(D)
Low	FM	126	432	.29
High	FM	121	432	.28
Low	CW	164	432	.38
High	CW	198	432	.46

It can be seen that p(D), for each condition, fell approximately in the range 0.3-0.5, indicating that, on the average, the signals were in the "difficult-to-detect" region which was desired for this experiment. Since the stimuli were designed to produce this level of performance, it should be obvious that these p(D) estimates cannot be regarded as evaluative of the operational performance of the AN/SQS-26 sonar, beyond the circular conclusion that about half of any signals at the "minimum detectable level" will be detected.

These p(D) estimates and the effects of the independent variables on them are presented graphically in the section on the "receiver operating characteristic" which follows the discussion of p(FR).

Probability of False Report, p(FR). An estimator of p(FR) may be formed from the ratio of the number of false reports to *the number of ways in which the non-signal stimuli could have elicited contact reports.* The latter term is easily defined for experiments in which detection opportunities occur as discrete events on a simple display. In the complex task of sonar detection, however, the number of false response opportunities presented depends on many factors. The principal factors are discussed in turn as they pertain to the sonar detection situation.

1. Characteristics of non-signal sources in the water which can give rise to contact reports: this factor is of obvious importance in considering false reports. For simulated noise/reverberation backgrounds, this factor is defined by the characteristics of the noise generating apparatus.

2. Volume of water searched, and the spatial resolution with which it is searched (as realized at the sonar display): this factor determines the number of opportunities the non-signal sources will have to mark the sonar display on a single ping (assuming an isotropic distribution of these sources; otherwise, their distribution must be taken into account). This factor may be specified in terms of the time-bandwidth product of the non-signal information presented on the entire display during one ping.

3. Characteristics of the operator as a detector: the first two factors deal with the detection situation *up to the sonar display*; the human, however, is an integral part of the detection process, and the human factor in this context is as important as the other factors (but, unfortunately, much less well defined) in formulating a model of the process. The "detection strategy" employed by the sonar operator certainly involves correlation of information across several pings, so that a single mark on the display can rarely be regarded as a "false response" opportunity. This strategy may be

> considered in the context of the usual signal detection model as reducing the number of opportunities calculated on the basis of individual pings, or it can be naturally accommodated in a sequential detection model.

The first two factors can, in principle, be accurately specified (for an experiment) from the electronic characteristics of the noise/reverberation source and the sonar set. The human factor mentioned above, however, has never been thoroughly investigated for complex displays, and is a ubiquitous source of speculation wherever the calculation of false response probabilities from experimental data are desired. No objective means exist to precisely define the effective number of false response opportunities presented in a man-machine detection system with complex displays, and there are nearly as many assumptions regarding this number as there are authors reporting experimental results. We intend to make another such assumption here; and we caution the reader who wishes to compare p(FR) *for various experiments* to carefully study each experimenter's "false response opportunity" assumptions.

We have assumed that there were 100 opportunities (per operator) for the noise/reverberation background in the 72 six-ping sequences to elicit false FM contact reports; and 100 opportunities for the 72 sequences to elicit false CW contact reports. We justify the *order of magnitude* of these assumptions in the following way: there were certainly more than 10 opportunities in the 72 sequences, since more than 10 false reports were made by some subjects during each session (for both displays); and, from our knowledge of the noise/reverberation characteristics and the detection behavior of the operators, we are convinced that there were fewer than 1000 opportunities for the non-signal information to stimulate contact reports. However, we do not intend to defend the number "100" over the number "50," or "200"; by the same token, we doubt that any other students of this situation will feel in a position to challenge the assumption by a factor this small.

Based on these assumptions, there were a total of 1800 false report opportunities over the 18 operators for each signal type and level of command attention. Thus, we may calculate p(FR):

COMMAND ATTENTION	SIGNAL TYPE	NUMBER FALSE REPORTS	OPPOR-TUNITIES	p(FR)
Low	FM	87	1800	.048
High	FM	43	1800	.024
Low	CW	284	1800	.158
High	CW	195	1800	.108

Receiver Operating Characteristic. Detection and false response probabilities are often jointly plotted in a format known as the "receiver operating characteristic" (ROC), with p(D) on the ordinate and p(FR) on the abscissa. The probabilities calculated in this section from the experimental data are displayed in the ROC format in Figure 20. The axes in this figure are not linear in "probability," but they are linear in the "unit normal deviates," or "standard scores," corresponding to the probabilities indicated on the axes. This property is useful because the usual signal detection theory model assumes Gaussian equal-variance distributions of the results of observations under noise and signal-plus-noise conditions. The outcome of this assumption is that in the "normal-normal" coordinate system of Figure 20, the locus of all operating points (i.e., pairs of probabilities) accessible to a "receiving system" with a specific detection efficiency, or "signal detectability," d', describes a *straight line* parallel to the "chance diagonal" shown in the figure, which is the locus of the points p(D) = p(FR). Each

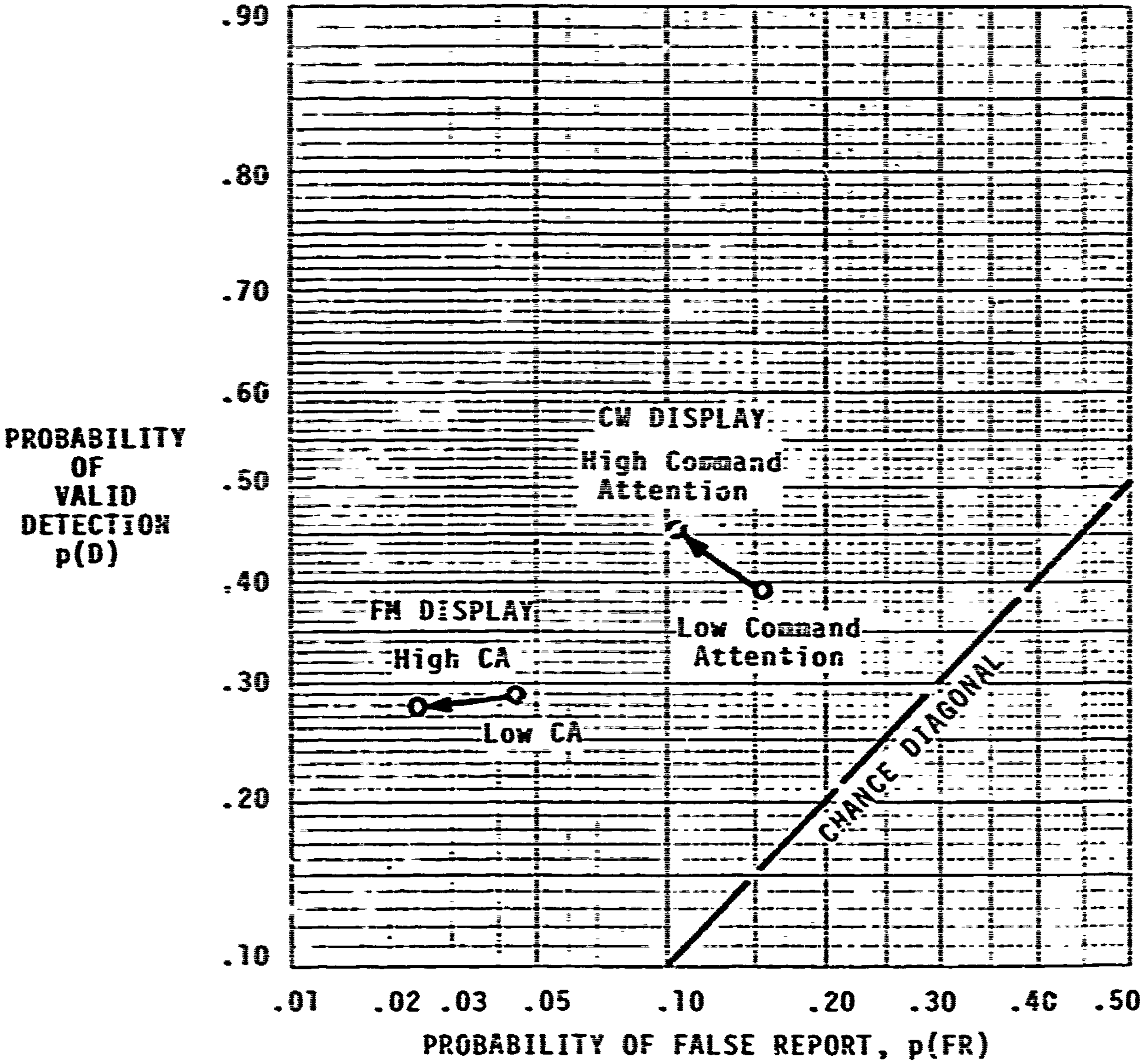


Figure 20. Receiver Operating Characteristic (ROC) showing experiment detection performance by signal type and level of command attention.

of the points on a line of constant signal detectability corresponds to a specific value of the detection criterion, β. Thus, according to this model, a change in detection criterion must move the system operating point *parallel* to the chance diagonal; and a change in effective signal detectability must move the system operating point *perpendicularly* to the chance diagonal. (See Green and Swets, 1966, for a detailed explanation of signal detection theory in psychophysics.)

It can be seen from Figure 20 that the "high" level of command attention moved the CW operating point away from the operating point for the "low" level *almost perpendicularly to the chance diagonal*, in a direction of improved effective signal detectability; and the displacement of the FM operating point had approximately equal components of perpendicular movement, corresponding to improved detectability, and parallel movement, corresponding to a change of reporting criterion in the direction of increased "caution." Thus, in the context of the usual signal detection theory model, the principal effect of the "high" level of command attention *was to improve system performance by improving detection efficiency*.

This conclusion, however, does *not* depend upon the assumption of the usual mathematical model of signal detection. *Any* reasonable model of the detection process would have to attribute improved detection efficiency to the effects of the "high" level of command attention upon the operating points shown in Figure 20, compared to the "low" level.

The Statistics "d'" and "β." In the signal detection model which assumes Gaussian equal-variance distributions of the results of observations, the measure of signal detectability is the separation of the means of the noise-only and signal-plus-noise distributions (along the axis of the "test statistic" random variable) in units of the common standard deviation, and is almost universally denoted by the symbol "d'." This parameter and the likelihood-ratio decision criterion, denoted by the symbol "β," together completely specify p(D) and p(FR), through the mathematical procedures specified by the model. Conversely, if an experiment is thought to be adequately described by this model, the experimentally observed p(D) and p(FR) may be used (through inverse mathematical procedures) to derive d' and β for that detection situation. *Alternatively, d' and β may be derived from experimentally observed p(D) and p(FR) through these procedures, to be used solely as descriptive statistics*, without an implied endorsement of the appropriateness of the underlying mathematical model. This empirical use of d' and β as descriptive statistics is both useful (in describing the joint effects of changes in detection and false response probabilities) and very widespread in the literature of detection experiments. It is in this latter sense that we calculate d' and β for this experiment, since, as we discussed in the first chapter, we doubt that the simple Gaussian equal-variance model is the most appropriate one for the sonar detection task.

A convenient table for calculating d' and β from experimentally observed proportions p(D) and p(FR) is given by Hochhaus (1972), and was employed to calculate the following values:

COMMAND ATTENTION	SIGNAL TYPE	d'	β
Low	FM	1.09	3.32
High	FM	1.47	6.95
Low	CW	0.69	1.57
High	CW	1.13	2.11

It can be seen that, on the average, signal detectability was greater for the FM signals than the CW; and that more "strict" reporting criteria (i.e., larger values of β) were employed for FM signals than for CW. It is also seen that the "high" level of command attention brought about a 35% improvement in FM signal detectability, and a 64% improvement in CW signal detectability, compared to performance at the "low" level. The reporting criterion concomitantly increased by 109% in the direction of "caution" for FM signals, and by 34% for CW signals.

Relationship of Attitude Toward "False Contact Consequence" to Numbers of False Reports Actually Made

In Chapter II we discussed the diversity with which officers and sonar operators rank ordered the tactical scenarios regarding false contact consequences. The questionnaire portion of that survey instrument and the rank ordering task with respect to false contact consequences were administered to the sonar technicians who served as subjects in the detection experiment, to permit comparison of their responses to those of the 219 men sampled in the survey. These tasks were administered at the end of each operator's last experimental session, via the instructions shown in Appendix G.

The responses of the operators to the questionnaire were not significantly different from those obtained in the main survey. They were also asked a supplemental question concerning the similarity of the experimental stimuli to the A-scan display of their own sonars. None of the operators indicated that the stimuli were significantly different from the display with which they were familiar.

The ranking task was administered to permit examination of the potential relationship between attitudes toward false contact consequences and actual false reporting behavior during the detection experiment. The behavioral criterion to be used in this comparison is obviously the number of false reports made by each subject. The attitudinal criterion to be compared with the number of false reports, however, is not so evident. The attitude data for each operator consisted of the rank order in which he placed the 17 scenarios in relation to the consequences of false alarms; we wished to derive a single descriptive statistic from these ranks, to be compared with the number of false reports that the operator made. This statistic was derived for each operator by comparing his rank ordering of the scenarios with the rank ordering shown in Figure 5, which was derived by scaling the rankings of the 119 officers and sonar technicians in Group I who judged false contact consequences to be worst for wartime scenarios.

The scenarios in Figure 5 have been assigned scale values, derived by the variability judgment technique discussed in Chapter 2. The responses of a single individual, however, cannot be assigned scale values, because the scaling technique depends upon variability in responses among *several* judges. Therefore, we could not use the scale values shown in Figure 5 in any comparison to individual rankings performed by the operators who participated in the experiment, which contain only ordinal information. Consequently, we determined the similarity of the *rank order* of the scenarios shown in Figure 5 with the rankings performed by *each* of the 18 operators in the detection experiment, by calculating Spearman's rank correlation coefficients.

Thus, we obtained an attitudinal criterion for each operator, which reflects a comparison of his attitudes toward false contact consequences with the "Group I" attitude (that false contact consequences are *worst* in wartime situations). If the rank correlation coefficient calculated for an operator were near +1.00, it would indicate that his rank ordering was very close to that shown in Figure 5, and that he was probably a "member" of Group I. If the rank correlation coefficient were near -1.00, it would indicate that his rank ordering was nearly *opposite* to that shown in Figure 5 (i.e., very close to that shown in Figure 6), and that he was probably a "member" of Group II, which judged false contact consequences to be *least* severe during war. If the correlation coefficient were near zero, it would indicate that there is little correspondence between his rank ordering and those of either Group I or Group II, placing him in the Group III attitude category.

The total number of false reports which each subject made during the "high" command attention condition is shown in Figure 21 plotted against the corresponding rank correlation coefficient. Each point in this figure thus characterizes an individual operator's false reporting behavior, and his "false contact consequence" attitudes.

It was found that amount of experience with the AN/SQS-26 sonar was an important factor in the distribution of points in Figure 21. Two operators informed us that their primary duties involved the maintenance of fire-control equipment, and that they had no experience with the AN/SQS-26 A-scan display. These operators made very few false *or* correct reports during the experiment, reflecting an atypically conservative reporting behavior which was no doubt the result of their unfamiliarity with the stimuli. Therefore we discount the data points in Figure 21 (denoted by the symbol "X") corresponding to these operators. (The inclusion of these operators was found to have negligible impact on the other results reported in this chapter, so this is the only analysis from which they were excluded.)

Of the remaining 16 operators (whose primary duties *did* involve operating the AN/SQS-26 sonar), 6 had less than a year of experience with the sonar, and 10 had a year or more of experience. Figure 21 shows that there was a tendency for the less experienced operators, whose data points are indicated by open circles, to make more false reports than the more experienced group, whose data points are indicated by closed circles.

The figure also reveals a tendency toward making *fewer* false reports by the operators who felt that wartime false contact consequences were *worst* (i.e., those operators toward the right end of the scale). The dotted line in the figure

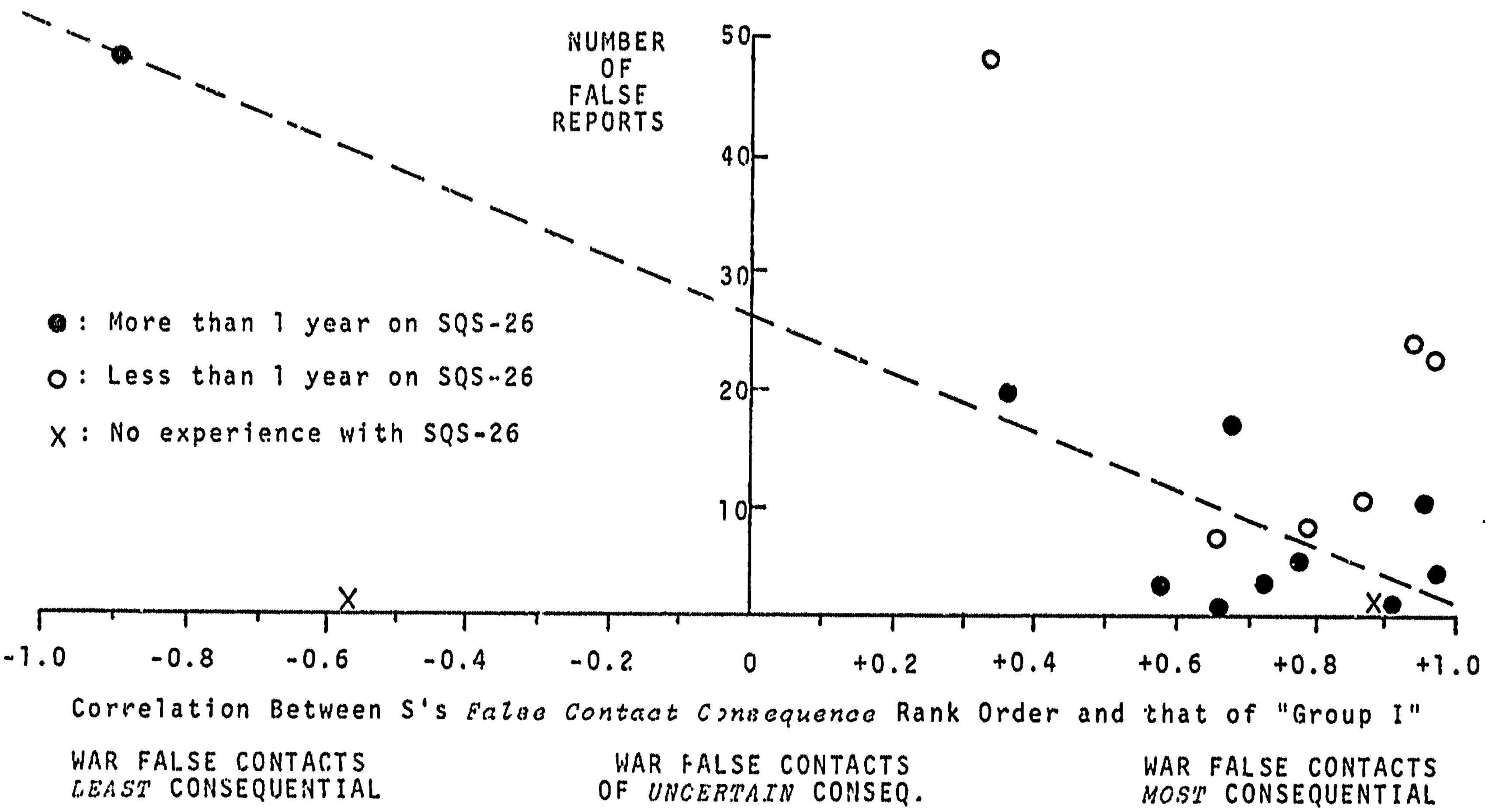


Figure 21. Number of false reports versus attitudes toward false contact consequences.

is a least-squares "best fit" to the data points for operators with one or more years of experience with the sonar, and the Pearson correlation coefficient between the numbers of false reports and the "attitude" statistics for this group was -0.92, which is a very significant value even for this small number of operators (p < .001).* The data points for the less experienced operators fall above this regression line (with one exception).

These data support the expectation that those operators who judged wartime false contact consequences to be *worst* would make *fewer* false reports during the "high" command attention condition. We cannot regard these data as conclusive, however, because of the limited sample size, and particularly because the sample (through the unfortunate medium of chance) clustered toward the right end of the "attitude" scale.

---

*It is recognized that one operator who evidently considered false alarms quite inconsequential contributed heavily to this correlation.

CHAPTER IV

CONCLUSIONS

Many detailed conclusions have already been drawn in the text of this report, in conjunction with the exposition of specific results. It is hoped that the reader will find in these specific results cause for further conclusions, and we hope that the many details which have been presented will facilitate that use of the research results.

However, in this chapter we wish to step back from the myriad technical details of psychometrics, statistics, signal detection theory, and so forth, to view the research in a more comprehensive perspective. In the first sentence of the first chapter of this report, we stated that "it is necessary for the Navy to make quantitative, detailed, and accurate estimates of expected performance in a wide range of threat situations, and particularly in accomplishing its potential wartime missions."

We hope to accurately interpret, in an integrated manner, the implications of this research concerning the expected performance of our destroyers in accomplishing their potential ASW missions.

Noise, Reverberation, and "False Contacts"

Most theoretical and laboratory studies of signal detection focus upon the task of differentiating a "signal-plus-noise" condition from a "noise-alone" condition; indeed, this task defines what is usually meant by "signal detection." An extensive theoretical framework has been built around this detection situation, and many experiments have been performed involving the human as the "detector," laying a theoretical and empirical foundation which has permitted us to conduct a meaningful sonar detection experiment, and analyze the sonar operator's detection behavior in response to operationally important variables.

However, an extremely important practical distinction exists between the "typical" laboratory detection task and the actual active sonar detection situation. In the "pure" detection task, performance is generally accompanied by "valid" detections, stimulated by the "signal-plus-noise" condition, and "false" detections, stimulated by the "noise-alone" condition. In any non-trivial detection task of this type, the "noise" background will be a significant source of false responses, and this potential for the noise to elicit detection responses is crucial to the underlying mathematical model of signal detection. In the sonar detection experiment described in Chapter III, all false responses *were* stimulated by the sonar noise/reverberation background, permitting the application of signal detection theory techniques to analyze the influence of the independent variables upon sonar operators' reporting behavior, the determination of which was the sole objective of the experiment.

However, in the operational sonar detection situation, the noise/reverberation background is an utterly insignificant source of "false" responses, compared with that unfortunately large class of objects which give rise to "false contacts." "False contacts" produce *real* echoes, whose characteristics differ significantly from the noise background (and, often, insignificantly from "valid" target echoes). The true noise and reverberation background is always uncorrelated over a time interval of several pings, and can therefore produce consistent marking of the sonar display at a given location only with vanishingly small probability; the "false contact," however, persists, and is reported. Thus, the noise/reverberation background as a source of false reports leading to seriously inappropriate tactical responses is one of the few *minor* problems which confront destroyer ASW.

We are, of course, talking about the classification problem. One cannot criticize signal *detection* theories or experiments for not dealing with classification (i.e., the task of distinguishing one signal from another, rather than signal-plus-noise from noise alone), since, by definition, it is not in their domain to do so. By the same token, however, it is almost by definition impossible to meaningfully discuss destroyer ASW *without* considering the classification problem.

> "Classification continues to be the most pressing problem in ASW." A Chief Sonar Technician whom we interviewed.

> "Classification continues to be the biggest problem." A Destroyer ASW Officer.

> "I believe rapid classification is still the major problem in command and control in ASW." A Destroyer Commanding Officer.

False Contacts: The High Threshold Solution

Without false contacts, of course, there would not be any classification problem. The fact is, however, that the oceans contain many more "things" that *look* like submarines, to a sonar, than there are submarines; and as we continue to trade cue resolution for longer ranges in our newer sonars, these "things" look more like submarines than ever before. The peacetime solution to this high incidence of false targets is simple and well-known:

> "When returning from a cruise, whether long or short, it is common practice for sonar to get the unofficial word 'NO CONTACTS.' If sonar should report a contact during this time, whether sub or not, sonarmen will not get a 'WELL DONE.' Hence, reporting during this period is almost non-existent, unless the submarine forces a report, by surfacing 1,000 yards away for example." A Chief Sonar Technician.

> "Although ASW is the primary mission of a destroyer, I have found that the present-day officers on the bridge have little or no interest in the prosecution of a sonar contact, be it possible, or non-sub, during any underway period, except an ASW exercise. ...Remarks like 'turn your gain down, and it will go away,' or 'we have something more important to do' [account for decreased contact reporting]." Another Chief Sonar Technician.

This attitude from the bridge has the effect of raising sonar's reporting threshold to a high value--they report contacts only when very confident. However, it is not true that all officers share this lack of interest in prosecuting contacts, or think the "high threshold" solution to false contacts is a good one:

> "In general, the practice has been to report only if you are sure of a contact (especially in a situation where you are being evaluated). I believe the sonar team must be encouraged to report immediately any contact, especially with the advent of the sophisticated sonar in use today. The need for immediate response precludes the accumulation of 'pings' in order to have a firm evaluation prior to reporting a contact to ship control. The requirement for sonar to report any and all contacts immediately must be stressed in order for the unit to take the necessary action to maintain and prosecute a contact." A Destroyer Executive Officer.
>
> "False contacts, in general, are not nearly the problem that fewer, or no contacts are. I feel that in most cases, low confidence contacts deserve more investigation than they customarily receive." A Destroyer Commanding Officer.
>
> "Aggressiveness cannot be sacrificed at the expense of legitimacy. I don't feel the false contact rate is so detrimental, as training and practice are gained by initial classification. Bridge watches should heed sonar more often. The bridge is too often concerned about 'sticking out its neck.' The only way contacts can be gained is by reporting and prosecuting." A Destroyer Operations Officer.

Nonetheless, it is very apparent even from these comments that the peacetime reporting threshold is a high threshold.

The scaling of 17 peacetime, ASW exercise, and wartime scenarios by our sample of 99 destroyer officers and 119 sonar technicians according to "certainty necessary for reporting" clearly indicates that much more confidence in a sonar contact is required to elicit a report during peacetime non-ASW operations (see Figure 3). We may also conclude, from the results of the detection experiment, that the typically "low" level of command attention focused upon peacetime sonar detection performance brings about a decrement in detection efficiency, or "effective signal detectability."

The consequence of this high threshold, low detection efficiency mode of operation is to screen out false (and valid) contacts at the lowest hierarchical level, relieving the rest of the team (CIC, U3, and command) of the burden *and the experience* of the ASW decision-making process. In time of war, however, that burden is going to shift.

The Wartime Threshold: "Report Everything"

> "When you're in a time of war--speaking for myself, I would be scared as hell and probably report more contacts...". An STG3 Sonar Operator.
>
> "Obviously, more false contacts will occur in wartime due to the fact that sonarmen will, and should be, less reluctant to call and classify a contact." A Destroyer ASW Officer.

"Will, and should be, less reluctant...": a predictive, and a normative statement. We agree with both. The scaling of scenarios with respect to "certainty necessary for reporting" (Figure 3), shows the wartime scenarios to be in a class by themselves--far removed from the other scenarios in the direction of a very "free" reporting threshold. Likewise, scaling of the same scenarios with respect to "missed contact consequences" (Figure 4) and with respect to "consequences of delayed contact reporting" (Figure 10) indicates consequence evaluations which will result in a very low reporting threshold.

There is remarkable agreement between officers and sonar operators (and junior/senior personnel, and LANTFLT/PACFLT groups) in their responses to our interviews, and they indicate much "freer" reporting in time of war.

We agree that the wartime reporting threshold *should* be low for the obvious reason that the detection/classification/attack sequence must be a group/hierarchical process with maximum information flow to the *top*, where, in time of war, the decision *must* be made. There is no way, theoretically or practically, to make optimal decisions by arbitrarily restricting information flow from the primary sensor, particularly in the information-starved ASW environment.

Of course, this is going to give rise to a considerable problem on the bridge.

The Decision Problem

Half the men in our survey sample who estimated wartime deepwater false contact rate expected more than 10 false contacts per day. What is the magnitude of the implied decision problem? The reader familiar with ASW exercise results may detect the ring of unrealistic optimism in this commanding officer's comment:

> "If equal action is assumed for each false contact, there would be a very serious problem. I am assuming that proper evaluation by the team will evaluate 90% or more of the false contacts as false contacts and that only a small number will result in firing or leaving the screen, etc." A Destroyer Commanding Officer.

As this destroyer executive officer puts it:

> "In a wartime situation, I would expect the old problem of expenditure of ordnance on false contacts to repeat itself."

The "old problem," however, is more of a problem than ever, because weapons are greater in cost and fewer in number, and because the fast nuclear submarine target forces the necessity of much quicker response upon us, with an attendant higher probability of attacking false targets, and lower probability of detecting and attacking real ones.

And to make the ASW decision problem more complex and unpredictable, we have found that the decision makers take widely divergent views concerning false contacts. In ranking the 17 scenarios according to consequences of false contacts, 55% of the sample ("Group I") were very much in agreement among themselves in ranking the wartime false contact consequences worst (Figure 5). On the other hand, 28% of the sample ("Group II") took a completely opposite view of false contact consequences, considering false contact consequences *least* severe during war (Figure 6). The Pearson correlation coefficient between the scenario scale values of Group I and Group II was *minus* 0.972--almost perfect mirror images. Then, there is Group III--a group comprising 17% of the sample, consisting of officers and operators in almost complete disagreement among themselves, and with respect to the other groups, concerning the consequences of false contacts.

It is likely that a person's views of false contact consequences will have an effect upon his decision-making behavior. We see an effect on our "contact reporting threshold" scale when Group I and Group II are independently plotted (Figure 8). Those who feel false contacts are least significant during war (i.e., Group II) are seen to advocate a greater change in "reporting threshold" from peace to war than do the members of Group I. Since each subject was required to rank the scenarios according to "reporting criterion"--i.e., confidence necessary to report--*before* he was required to rank them with respect to false contact consequences, he may not have explicitly taken into account his feelings toward false contacts while responding on the former scale. That is, the effect of the Group I/Group II differences on the "reporting threshold" scale may be conservative.

We also see a relationship between the false contact consequence attitudes of the sonar operators who participated in the detection experiment, and the numbers of false reports which they made (Figure 21). It is also probable that attitudes will influence officers' decision-making performance.

Analysis of the survey data shows no way to predict subjects' membership in Group I, II, or III from the personal data collected. No group had a membership which was characteristically officer, enlisted, LANTFLT, PACFLT, "more experienced," "less experienced," or composed of commanding officers. It follows that whatever influences the attitudes of Groups I, II, and III have on decision-making behavior in ASW are now distributed unpredictably. (It is probable, however, that these variances can be minimized through training and procedural implementations.)

In summary, it seems likely that at the onset of war, the naturally large number of objects which can cause false contacts, the greater wartime inclination for the sonarmen to detect and report them, and the inexperience of the rest of the team in dealing with them (and their heterogeneous attitudes toward them) will constitute a very serious ASW decision problem.

The Equilibrium State

We have concluded that the peacetime destroyer ASW posture is often one of low command attention, low detection efficiency, and of comparatively "reluctant" sonar contact reporting behavior; and we have concluded that at the onset of an ASW war, the pendulum will swing to the opposite extreme. The latter condition, however, will necessarily be a transient state. Sonar contact detection, reporting, prosecution, and attack cannot long continue in a rampant manner; we will not be able to "classify" every contact with a weapon. Historical evidence indicates that after some period of time, a tolerable ratio of false to valid attacks will be achieved (Buchanan and Freilich, 1971). But, how will this "equilibrium state" be effected?

We fear that unless destroyer officer personnel appreciate the ASW problem in its fullest extent, particularly including the sonar operator's very difficult task, apparent command dissatisfaction with the "transient state" detection and classification performance will, intentionally or unintentionally, influence the sonar operator toward undesirably conservative detection behavior. And it is a discouraging fact that very many destroyer commanding officers have been provided with very little ASW training or practical experience.

The equilibrium state must not be achieved by the strangulation of sensor information at the sonar operator's level, inadvertently or otherwise; it will ideally be achieved on board each destroyer through the informed decisions of a competent, professional ASW *team*. In order to attain this ideal, it seems that we must begin to treat surface ship ASW as the demanding, highly technical, highly specialized discipline that it is, and begin to develop significant numbers of Naval officers who have been given the thorough training which that discipline demands, as many of the other important navies have already done. A detailed diagnosis and prescription regarding this matter is given by Mackie (1972).

In the meantime, the research reported here leaves us with questions, as research usually does. How will the decision-making behavior of *today's* destroyer officer be influenced at the onset of ASW war, and how, in turn, will he influence the sonar operator's decision behavior, in passing to the equilibrium state? We are presently seeking answers to those questions.

REFERENCES

Birdsall, T. G., & Roberts, R. A. Theory of signal detectability: deferred-decision theory. *Journal of the Acoustical Society of America*, 1965, *37*, 1064-1074.

Broadbent, D. E. *Decision and Stress*. New York: Academic Press, 1971.

Buchanan, W. B., & Freilich, E. D. The False Attack Question in ASW (U). *Center for Naval Analysis, Operations Evaluation Group*, Rep. No. CRC 191, 1971, (SECRET).

Green, D. M., & Swets, J. A. *Signal Detection Theory and Psychophysics*. New York: John Wiley and Sons, 1966.

Harabedian, A., *et al.* Detection and Classification Performance of Airborne Lofar Operators: An Empirical Evaluation of Performance Models (U). *Human Factors Research, Inc., Tech. Rep. 785*, 1970, (SECRET).

Hochhaus, L. A table for the calculation of d' and β. *Psychological Bulletin*, 1972, *5*, 375-376.

Jeffress, L. A. Contributions of Psychophysics to Sonar. *Applied Research Laboratories, The University of Texas at Austin*, Rep. No. ARL-TM-69-23, 1969.

Kostoff, M. R., & Montgomery, M. B. Summary Technical Report, Volume II. Effect of Mutual Interference on AN/SQS-26 (CX) Sonar Performance as Determined in Human Factors Experiments Using Simulated Sonar Displays (U). *TRACOR*, Doc. No. T70-AU-7385-C, 1970, (CONFIDENTIAL).

Kruskal, J. B. Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. *Psychometrika*, 1964a, *29*, 1-28.

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964b, *29*, 115-130.

Mackie, R. R. The ASW officer: "Jack of all trades, master of none." *Proceedings of the United States Naval Institute*, 1972, *98*, 34-40.

Mosteller, F. Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. *Psychometrika*, 1951, *16*, 3-9.

Rizy, E. F. Effects of Decision Parameters on a Detection/Localization Paradigm Quantifying Sonar Operator Performance. *Raytheon Company, Submarine Signal Division,* Rep. No. R1156, 1972.

Shepard, R. N. The analysis of proximities: multidimensional scaling with an unknown distance function: I. *Psychometrika.* 1962a, *27,* 125-40.

Shepard, R. N. The analysis of proximities: multidimensional scaling with an unknown distance function: II. *Psychometrika,* 1962b, *27,* 219-246.

Stern, H. W. A Survey of Procedures Used in the Operation of the AN/SQS-26 Sonar System (U). *Naval Personnel and Training Research Laboratory,* Rep. No. SRR 71-21, 1971, (CONFIDENTIAL).

Thurstone, L. L. A law of comparative judgement. *Psychological Review,* 1927, *34,* 273-286.

Wald, A. *Sequential Analysis.* New York: John Wiley and Sons, 1947.

Williges, R. C. The role of payoffs and signal ratios in criterion changes during a monitoring task. *Human Factors,* 1971, *13,* 261-267.

Winer, B. J. *Statistical Principles in Experimental Design.* New York: McGraw-Hill Book Company, Inc., 1962.

Young, F. W., & Torgerson, W. S. TORSCA, a FORTRAN IV program for Shepard-Kruskal multidimensional scaling analysis. *Behavioral Science,* 1967, *12,* 498.

APPENDIX A

OPERATIONAL SCENARIOS

1. REFRESHER TRAINING

Your ship is undergoing refresher training at GITMO just after having completed a three-month shipyard overhaul. There are no other ships in company. Your ship is conducting damage control and engineering exercises.

2. UNIDENT

You are operating in the Sea of Japan (or in the Med.) in company with an ASW group. Ships in company include a CVS and 8 destroyers. One of the fixed-wing aircraft from the carrier has reported a disappearing radar contact 20 miles ahead of the task group. Helicopters have been dispatched from the aircraft carrier and are presently in the contact area. The helos are reporting intermittent sonar contact. Your ship and one other destroyer have been detached to proceed to the contact area.

3. WARTIME MERCHANT CONVOY

Your ship is one of six destroyers screening for a 50-ship merchant convoy steaming in a nuclear defense disposition, transiting from Halifax, Nova Scotia to Glasgow, Scotland. A state of war has existed between the NATO countries and the Soviet alliance for a period of one month. The convoy you are with is the third convoy to make the west/east North Atlantic crossing since the beginning of the war. The first convoy made the transit without incident. The second convoy encountered light opposition, losing four merchant ships. Intelligence reports indicate that several Soviet submarines are operating in the North Atlantic. Some of these submarines are known to be equipped with cruise (anti-shipping) missiles. The range of these missiles is about 30 miles.

4. REFRESHER TRAINING

You are undergoing ASW refresher training in the GITMO op area. You have been operating with a submarine for the past several hours but have not held contact for some time, an hour or more.

5. DEPLOYMENT TRANSIT

You are heading overseas at the beginning of what will be a 7-month deployment. Ships in company include an aircraft carrier, an oiler, and three other destroyers. There are no known submarines operating in the area.

6. ANNUAL COMPETITIVE EXERCISE

You are conducting your annual competitive exercise for the ASW-E. You are in company with two other destroyers, two helicopters, and two fixed-wing aircraft. At the beginning of the exercise, there is a submarine positioned at a point 12 miles ahead of the destroyers. At COMEX, the submarine submerged and is making an approach on the destroyers. In order to pass the exercise, the destroyers must detect and successfully attack the submarine.

7. OPPOSED SORTIE EXERCISE

You are getting underway from your home port to participate in a major fleet exercise. You know that your sortie will be opposed by at least two submarines. Your ship's assignment is to clear the harbor, proceed to a previously assigned patrol area, and to search that area until the other exercise units have cleared the harbor. Once the other units have cleared the harbor, you will form a screen around them and proceed to sea. Your job is to prevent the "enemy" submarine from obtaining a firing position on the screened units.

8. STEAMING AT NIGHT IN A LOCAL OP AREA

You are standing the midwatch on a ship steaming independently in a local op area. You have spent the week at sea conducting gunnery and seamanship exercises and will be conducting a gunnery exercise in the morning prior to returning to port.

9. WARTIME STRIKE OPERATIONS

A state of war exists between the U. S. and Communist China. You are steaming in company with a fast carrier strike group operating in EMCON, Condition 2 (no electronic emissions other than sonar). The mission of your task group is to conduct a surprise raid on China's coastal defenses. No enemy activity has been reported on the task group's ECM equipment and airborne early warning aircraft operating in advance of the task group have not reported any enemy activity. The probability is that your task group has not yet been detected by the enemy. Your task group is now in a position 200 miles east of the northern tip of Luzon, proceeding toward the Chinese mainland. There have been no reports of enemy submarine activity in this part of the Pacific for the past several days.

10. UNIDENT

You are in port overseas when word is received that an unidentified submarine has been sighted operating in the vicinity. Your ship and one other destroyer get underway to investigate. The original sighting was made by a U. S. patrol aircraft and the DATUM is considered accurate to within 2 miles. The contact was held by the aircraft for 5 minutes and then lost and not regained. You arrive on the scene 30 miles from the port 4 hours after the contact was originally reported.

11. WARTIME SCREENING FOR AN AMPHIBIOUS TASK GROUP

Your ship is one of 12 destroyers screening for a fast amphibious task group. In addition to the amphibious forces present, your movement group is being supported by a CVA and a CVS. The CVA and CVS are operating in a support position 50 miles on the flank of the amphibious group. The mission of your task force is to effect a landing on mainland China. Although a state of war has existed between the U. S. and Red China for a period of 2 months, the Soviet Union has remained neutral. However, the Soviet Union has threatened to enter the war if the U. S. should attempt a landing on mainland China. Present position of your task group is 200 miles east of Okinawa headed toward mainland China.

12. STEAMING IN A LOCAL OP AREA, ACTING AS A NIGHT PLANE GUARD DURING CARRIER AIR OPERATIONS

Ships in company are the aircraft carrier and one other destroyer. There are no known submarines operating in the area.

13. CONDITION III, STEAMING OVERSEAS

You are steaming in Condition III in the Mediterranean. You are in company with a CVA, and three other destroyers. There have been several recent reports of UNIDENTS in the area during the last month, and at least two Soviet submarines are known to be operating somewhere in the Med. at the present time.

14. SPECIAL ASW EXERCISE

You are operating in a large mid-ocean area in company with a task group consisting of an attack aircraft carrier and 3 other destroyers. There is also an "enemy" submarine participating in the exercise. The general nature of the exercise is to gather data on the ability of the screening ships to prevent the submarine from making a screen penetration and obtaining a firing position on the carrier. You know that the submarine will attempt to penetrate the screen sometime during the next 4 hours.

15. ASW TYPE TRAINING

You are conducting ASW type training in a local op area. Other units in company include a destroyer, two fixed-wing aircraft, and two helicopters. At the beginning of the exercise, a submarine was positioned on the surface 10 miles ahead. At COMEX, the submarine submerged. You are now attempting to close the target, gain contact, and conduct an attack on him.

16. WARTIME HUNTER KILLER OPERATIONS

Your destroyer is one of 8 destroyers operating with a CVS as a hunter killer group. Your mission is to operate in the general area of North Atlantic shipping lanes and to seek out and destroy enemy submarines. Although your task group has been credited with sinking 4 enemy submarines in the last 5 days, all of these sinkings have been the result of the action of the air group embarked on the carrier. None of the destroyers in company have reported any sonar contacts.

17. DEPLOYMENT TRANSIT

You are on your way *home* from a seven-month deployment in company with a carrier, an oiler, and three other destroyers. There are no known submarines operating in the area.

APPENDIX B

SONAR OPERATORS' AND OFFICERS' SURVEY INSTRUCTIONS

(Only those pages of the Officers' instructions which differ from the Operators' are included)

INTRODUCTION

FACTORS IN SONAR TARGET DETECTION/REPORTING

As you no doubt know, the Chief of Naval Operations regards ASW readiness/effectiveness information as highly important, and considerable effort has been expended to obtain such information through ASW exercises. Therefore, CNO/DOD planners also wish to know how such exercise data can be best related to the expected performance of our ASW forces during non-exercise situations. Human Factors Research, Inc., is under contract to the Navy to investigate certain aspects of this question. In order to do this, we wish to obtain your judgments (and those of several hundred other fleet ASW personnel) concerning various operational situations.

To obtain the desired information, we will ask you to *rank order* a set of scenarios (printed on cards) according to the first instruction sheet. To do this, you can spread the cards out in front of you and arrange them in a single column--but you will probably find it more convenient to actually form three shorter columns side-by-side, so you can easily see all of the cards. When you are satisfied with a particular ordering of the cards, please write the card numbers in the boxes on the reverse side of the instruction sheet, *and ask for the next instruction sheet.*

Please feel free to perform these rankings in a way which will reflect *your own opinion.* Our report to the Navy will be in terms of trends and averages; only our company research staff will see the actual questionnaires and answer sheets.

Before you begin the first ranking, please complete the biographical questionnaire on the other side of this page.

BIOGRAPHICAL DATA

Rank/Rate__________Length of Service__________Yrs.__________Mos.

Present Billet/Responsibilities________________________________

______________________________How Long?________Yrs.________Mos.

Educational Level--Highest Grade in School:

7 8 9 10 11 College Degree? [ ] Yes [ ] No

12 13 14 15 16 Graduate Study? _____________Yrs.

Navy Schools Related to ASW____________________________________

____

Area of Specialization in the Navy_____________________________

____

____

Previous ASW Billets____

____

____

____

Name of Ship____

Type of Ship____

(SONAR OPERATOR'S VERSION)

SONAR CONTACT REPORTING

For the purposes of this part, please imagine that you are standing a sonar watch aboard your ship, that you are directly operating and observing your sonar, and that *you and you alone* will make the decision to report sonar detections to the bridge. Your certainty of a contact depends on many things, such as echo quality, consistency, strength, and so on, and you can be more sure of some contacts than others. If you were to see/hear a very strong echo which showed obvious submarine target cues or characteristics, you could report "possible sub" to the bridge with little doubt or uncertainty concerning the contact. On the other hand, if you were to see/hear an "echo" which was very weak, inconsistent, and lacking in cues, you might not be sure that you actually have a contact. Your decision to report such a questionable "contact" to the bridge might depend upon the operational situation--for example, you might be more likely to report such contacts during wartime ASW operations than you would in non-ASW peacetime situations. Please carefully read and consider each of the situations described on the cards, and place the cards in rank order in front of you so that the situation in which *YOU FEEL YOU WOULDN'T NEED TO BE VERY SURE OF A CONTACT TO REPORT "POSSIBLE SUB" TO THE BRIDGE* is at the top, downward situation-by-situation, to the situation for which *YOU WOULD WANT TO BE PRETTY SURE OF A CONTACT TO REPORT.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

SONAR CONTACT REPORTING

*"I WOULDN'T HAVE TO BE VERY
SURE OF A CONTACT TO REPORT"*

☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐

*"I WOULD WANT TO BE
PRETTY SURE OF A CONTACT
TO REPORT"*

(OFFICER'S VERSION)

SONAR CONTACT REPORTING

For the purposes of this part, please imagine that you are at sea. Sonar's certainty of a contact depends on many things, such as echo quality, consistency, strength, and so on, and they can be more sure of some contacts than others. If they were to see/hear a very strong echo which showed obvious submarine target cues or characteristics, they could report "possible sub" to the bridge with little doubt or uncertainty concerning the contact. On the other hand, if they were to see/hear an "echo" which was very weak, inconsistent, and lacking in cues, they might not be sure they actually have a contact. You may feel that their decision to report such a questionable "contact" to the bridge could depend upon the operational situation--for example, they might be more likely to report such contacts during wartime ASW operations than they would in non-ASW peacetime situations. Please carefully read and consider each of the situations described on the cards, and place the cards in rank order in front of you according to your opinion of how sonar *should* report, so that the situation in which *YOU FEEL SONAR NEED NOT BE VERY SURE OF A CONTACT TO REPORT "POSSIBLE SUB" TO THE BRIDGE* is at the top, downward situation-by-situation, to the situation for which *YOU FEEL THEY SHOULD BE PRETTY SURE OF A CONTACT TO REPORT.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

SONAR CONTACT REPORTING

THEY NEED NOT BE VERY SURE
OF A CONTACT TO REPORT

☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐

THEY SHOULD BE PRETTY SURE
OF A CONTACT TO REPORT

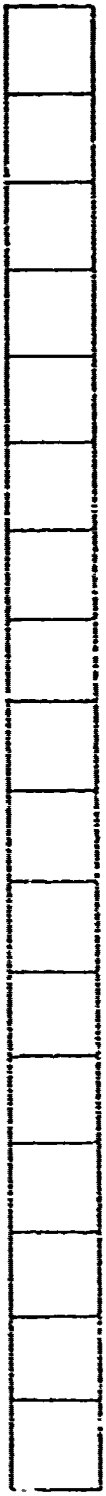
(SONAR OPERATOR'S VERSION)

CONSEQUENCES OF MISSED CONTACTS

For the purposes of this part, please imagine that you are the sonar operator and that a "contact" briefly caught your attention, but was so weak, intermittent, and lacking in cues that you had little confidence that it actually *was* a contact and did not report it to the bridge as "possible submarine." At least part of the time, such contacts could actually be caused by submarines. If it *were* actually a submarine, it is a missed contact situation, which can have a variety of consequences. Your ship may lose points during an exercise, in time of war your ship or those in company may be torpedoed, the submarine may move out of range and never reappear, with its existence remaining unknown, and so on. The exact consequences of a missed contact may be different for different people, and may depend upon the particular operational situation. *Please imagine that you have actually missed a contact.* Carefully consider what the consequences of a missed contact might be *as you see them* for each of the operational situations described on the cards, and place the cards in rank order in front of you so that the situation whose missed contact consequences are *MOST SEVERE, MOST UNDESIRABLE,* and/or *MOST UNPLEASANT* to you is at the top, downward situation-by-situation so that whose missed contact consequences are *LEAST SEVERE, LEAST UNDESIRABLE,* and/or *LEAST UNPLEASANT.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

CONSEQUENCES OF MISSED CONTACTS

CONSEQUENCES MOST SEVERE



CONSEQUENCES LEAST SEVERE

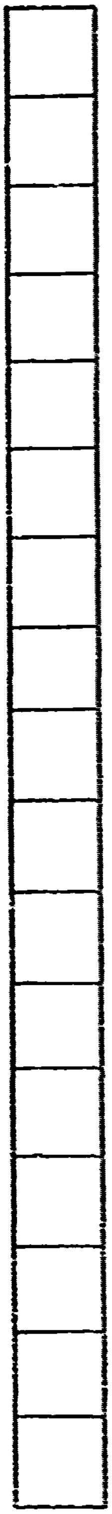
(OFFICER'S VERSION)

CONSEQUENCES OF MISSED CONTACTS

For the purposes of this part, please imagine that you are at sea and that a "contact" briefly caught the attention of sonar, but was so weak, intermittent, and lacking in cues that they had little confidence that it actually was a contact and did not report it to the bridge as "possible submarine." At least part of the time, such contacts could actually be caused by submarines. If it *were* actually a submarine, it is a missed contact situation, which can have a variety of consequences. Your ship may lose points during an exercise, in time of war your ship or those in company may be torpedoed, the submarine may move out of range and never reappear, with its existence remaining unknown, and so on. The exact consequences of a missed contact may be different for different people, and may depend upon the particular operational situation. *Please imagine that sonar has actually missed a contact.* Carefully consider what the consequences of a missed contact might be *as you see them* for each of the operational situations described on the cards, and place the cards in rank order in front of you so that the situation whose missed contact consequences are *MOST SEVERE, MOST UNDESIRABLE,* and/or *MOST UNPLEASANT* to you is at the top, downward situation-by-situation to that whose missed contact consequences are *LEAST SEVERE, LEAST UNDESIRABLE,* and/or *LEAST UNPLEASANT.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

CONSEQUENCES OF MISSED CONTACTS

CONSEQUENCES
MOST SEVERE



CONSEQUENCES
LEAST SEVERE

(SONAR OPERATOR'S VERSION)

CONSEQUENCES OF FALSE CONTACTS

For the purposes of this part, please imagine that you are the sonar operator and that you have reported a contact to the bridge as "possible submarine." At least part of the time, a contact reported as "possible sub" turns out to be non-submarine. This situation constitutes a false contact, and can have a variety of consequences. Fuel and/or weapons may be expended, the ship may leave a position in a screen unguarded, the ship's captain may have to be awakened, and so on. The exact consequences of a false contact may be different for different people, and may depend upon the particular operational situation. *Please imagine that you have actually reported a false contact.* Carefully consider what the consequences of reporting a false contact might be *as you see them* for each of the operational situations described on the cards, and place the cards in rank order in front of you so that the situation whose *FALSE CONTACT CONSEQUENCES WOULD BE MOST SEVERE, MOST UNDESIRABLE,* and/or *MOST UNPLEASANT* is at the top, downward situation-by-situation to that whose *FALSE CONTACT CONSEQUENCES WOULD BE LEAST SEVERE, LEAST UNDESIRABLE,* and/or *LEAST UNPLEASANT.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

CONSEQUENCES OF FALSE CONTACTS

CONSEQUENCES
MOST SEVERE

CONSEQUENCES
LEAST SEVERE

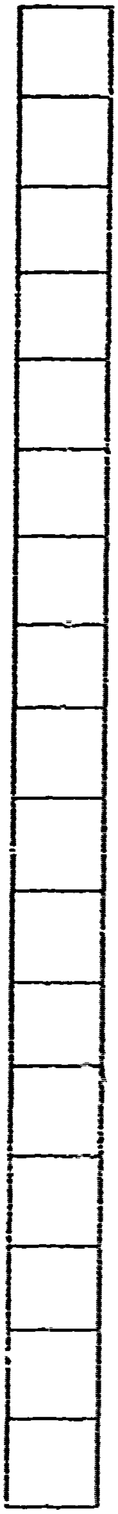
(OFFICER'S VERSION)

CONSEQUENCES OF FALSE CONTACTS

For the purposes of this part, please imagine that you are at sea and that sonar has reported a contact to the bridge as "possible submarine." At least part of the time, a contact reported as "possible sub" turns out to be non-submarine. This situation constitutes a false contact, and can have a variety of consequences. Fuel and/or weapons may be expended, the ship may leave a position in a screen unguarded, the ship's captain may have to be awakened, and so on. The exact consequences of a false contact may be different for different people, and may depend upon the particular operational situation. *Please imagine that sonar has actually reported a false contact.* Carefully consider what the consequences of a false contact might be *as you see them* for each of the operational situations described on the cards, and place the cards in rank order in front of you so that the situation whose *FALSE CONTACT CONSEQUENCES WOULD BE MOST SEVERE, MOST UNDESIRABLE*, and/or *MOST UNPLEASANT* is at the top, downward situation-by-situation to that whose *FALSE CONTACT CONSEQUENCES WOULD BE LEAST SEVERE, LEAST UNDESIRABLE*, and/or *LEAST UNPLEASANT*. After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

CONSEQUENCES OF FALSE CONTACTS

CONSEQUENCES
MOST SEVERE

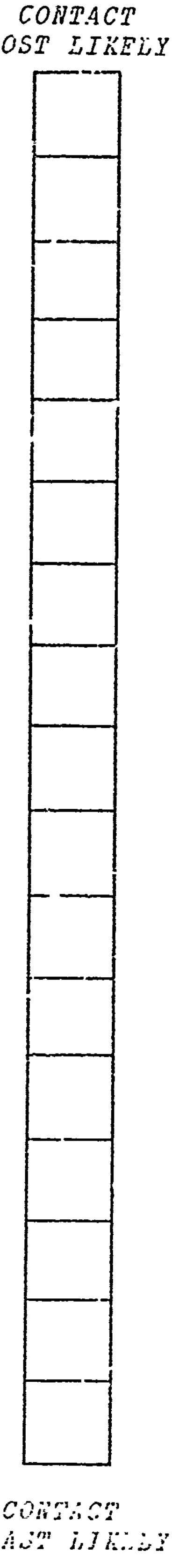


CONSEQUENCES
LEAST SEVERE

LIKELIHOOD OF MAKING CONTACT

The likelihood of actually making contact with a submarine varies from situation-to-situation. Please carefully consider each of the situations described on the cards, and place the cards in rank order in front of you so that the situation you judge *MOST LIKELY TO RESULT IN SUBMARINE CONTACT DURING ONE SONAR WATCH* is at the top, downward situation-by-situation to that *LEAST LIKELY TO RESULT IN SUBMARINE CONTACT DURING ONE SONAR WATCH*. After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

LIKELIHOOD OF MAKING CONTACT



RESPONSE TIME

A quick detection and report to the bridge may be more important in some situations than in others. Please carefully consider each of the situations described on the cards and place the cards in rank order in front of you so that the situation for which *DELAYING CONTACT REPORTING WOULD BE WORST* is at the top, downward situation-by-situation to that for which *DELAYING CONTACT REPORTING WOULD BE LEAST BAD*. After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

RESPONSE TIME

TO DELAY
IS WORST

☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐
☐

NO DELAY
IS LEAST
BAD

QUESTIONNAIRE

FOR EACH QUESTION, CHECK THE ONE RESPONSE THAT MOST ACCURATELY REFLECTS YOUR OWN OPINION.

1. During *present-day non-ASW peacetime operations* (for example, peacetime transiting, non-ASW refresher training, and so on), which do you feel sonar operators *should* do?

    ____ a. Report fewer false contacts than they do now even though it means increased delay in reporting.

    ____ b. Continue present reporting practices.

    ____ c. Report contacts more quickly than they do now even though it means more false contacts.

2. During *present-day ASW exercise situations*, which do you feel sonar operators *should* do?

    ____ a. Report fewer false contacts than they do now even though it means increased delay in reporting.

    ____ b. Continue present reporting practices.

    ____ c. Report contacts more quickly than they do now even though it means more false contacts.

3. Which have you found to be true?

    ____ a. Many more false contacts are reported during ASW exercises than during non-ASW peacetime operations.

    ____ b. A few more false contacts are reported during ASW exercises than during non-ASW peacetime operations.

    ____ c. About the same number of false contacts are reported during ASW exercises and non-ASW peacetime operations.

    ____ d. Fewer false contacts are reported during ASW exercises than during non-ASW peacetime operations.

4. How much of a problem do you think false contacts would be in the event of a war involving ASW?

    ____ a. Very serious problem.

    ____ b. Serious problem.

    ____ c. Moderate problem.

    ____ d. Insignificant problem.

5. At the present time, how well do you think the bridge can judge the degree of certainty which sonar has about an initial contact?

____ a. Very accurately.
____ b. Accurately.
____ c. Not very accurately.
____ d. Not at all.

6. How useful do you think knowledge of sonar's confidence or certainty concerning an initial contact would be in decision making on the bridge during a war involving ASW?

____ a. Very useful.
____ b. Moderately useful.
____ c. Of little use.
____ d. Of no use.

7. From *your own experience* during present-day ASW exercises, what do you feel is a typical false contact rate

a. in coastal/shallow water?

________________ per ______________
number of false contacts *unit of time*

b. in mid-ocean/deep water?

________________ per ______________
number of false contacts *unit of time*

8. *In the event of war* involving ASW, what do you feel a typical false contact rate might be

a. in coastal/shallow water?

________________ per ______________
number of false contacts *unit of time*

b. in mid-ocean/deep water?

________________ per ______________
number of false contacts *unit of time*

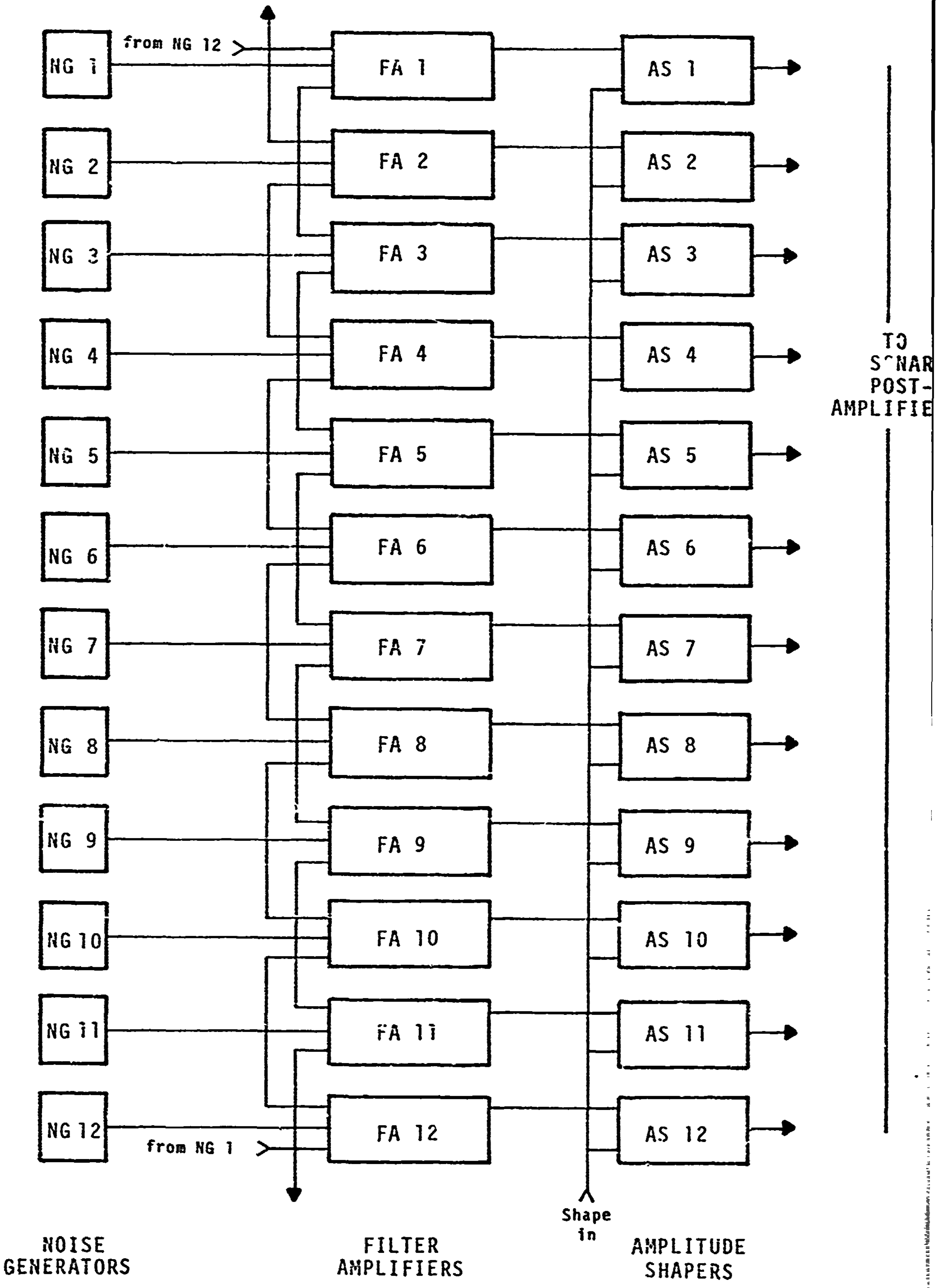
9. COMMENTS: ____

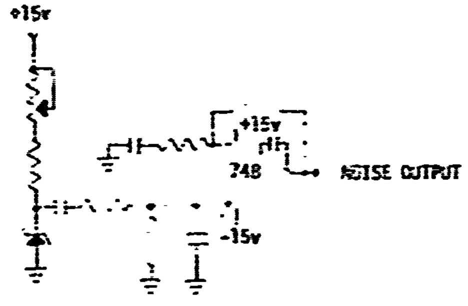
____

APPENDIX C

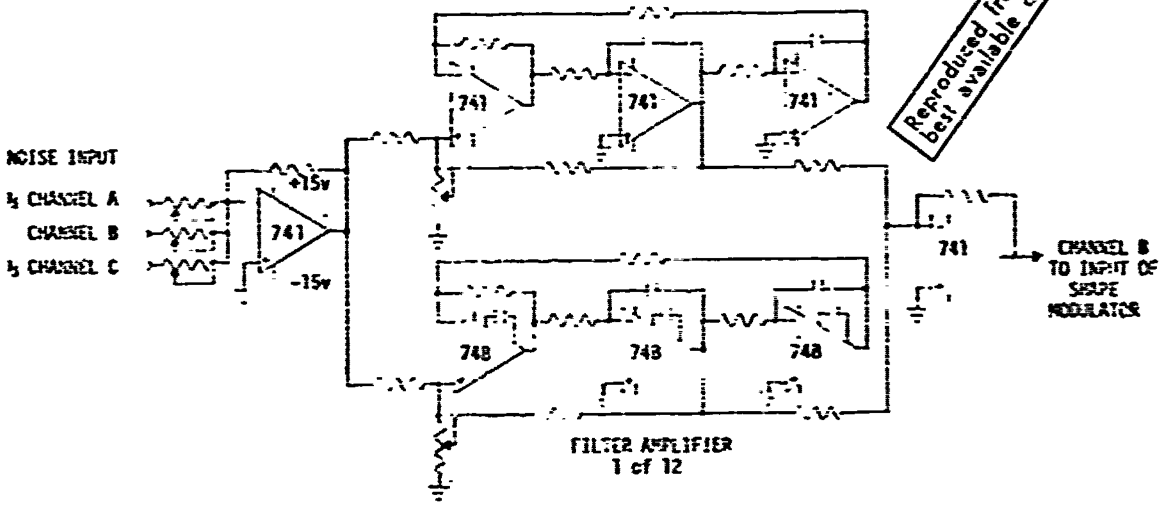
NOISE/REVERBERATION BACKGROUND SIMULATOR

NOISE/REVERBERATION GENERATOR BLOCK DIAGRAM

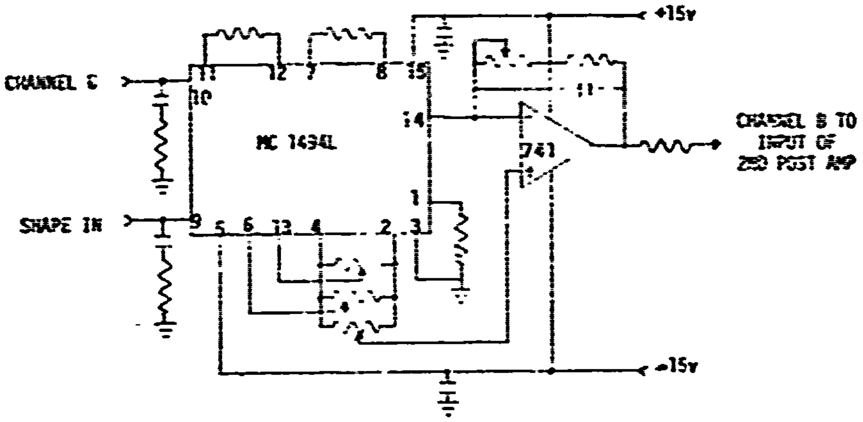




NOISE GENERATOR
1 of 12



FILTER AMPLIFIER
1 of 12



AMPLITUDE SHAPER
1 of 12

NOISE/REVERBERATION GENERATOR SCHEMATIC DIAGRAMS

APPENDIX D

INSTRUCTIONS USED IN THE SONAR DETECTION EXPERIMENT

SONAR CONTACT EVALUATION

You are being asked to view a number of photographic slides of an AN/SQS-26CX A-scan display to judge whether any signals appeared on the display that you would report as "possible sub" contacts during routine non-ASW operations at sea. The reason for doing this is that while the electronic operation of the SQS-26 sonar is well understood from the transducer to the displays, very little is known (outside the fleet) about what constitutes a reportable contact with this sonar during routine operations under actual sonar conditions. We do not know exactly how many contacts are in these slides, or how many would be reported. Therefore, we want your best judgment of what you would report to the bridge on your ship as "possible sub" contacts during routine non-ASW operations at sea.

Please feel free to report in a way which reflects the way you would report aboard your ship. You cannot be scored "right" or "wrong"; the results will simply be the average judgments of many operators from several ships regarding what they see in the slides that they think would be reported.

Slide Viewing Procedure

As you know, the SQS-26CX A-scan display has a 6-ping history. In the "single-storage" mode, and after a display erase, the outputs of the FM and CW processing for each ping are recorded side-by-side on the display until the display contains the results of 6 pings. All the slides you will be shown are grouped into "sequences" of 6 slides, which show the results of 6 consecutive pings. Prior to photographing each sequence, the A-scan display was erased, and the first slide of each sequence shows the results of one ping only. The second slide of each sequence was taken after the next ping, and because of the display memory the second slide shows the results of the first ping *and* the second ping. The third slide shows the results of pings 1, 2, and 3; the fourth slide shows pings 1, 2, 3, and 4; the fifth slide shows pings 1, 2, 3, 4, and 5; and the sixth (last) slide of a sequence shows the results of pings 1, 2, 3. 4, 5, and 6. Each and every 6-slide sequence progressively shows the results of 6 consecutive sonar pings which occurred one right after the other. *However,* the various *sequences* were obtained at different times, and there may be no relationship between what you see in one sequence and what you see in another. But *within* any given sequence, the ping history develops just as it did on the sonar display.

Your slide projector will automatically show you the "next" slide every 25 seconds. Therefore, you will be able to look at each slide for 25 seconds before seeing the next; that is, you will be able to look at the results of a given ping (and any prior pings which were stored on the display) for 25 seconds before seeing the results of the *next* ping. This is somewhat faster than you would be able to view successive pings at the sonar (the A-zone start was always at some range from own-ship, and you will be looking at a 20 Kyd zone width) but you should have ample time to scan the display after every ping.

Each slide you will see has a four-digit number in the lower right-hand corner which will permit you to tell which pings of which sequence it shows. The first two digits identify the "sequence"; the last two digits identify the "ping number." As you view successive slides, you will see the numbers develop in the following way:

SLIDE NUMBERS	WHAT YOU SEE IN THE SLIDE
0101	sequence 1, ping 1
0102	sequence 1, pings 1 & 2
0103	sequence 1, pings 1, 2, & 3
0104	sequence 1, pings 1, 2, 3, & 4

SLIDE NUMBERS	WHAT YOU SEE IN THE SLIDE
0105	sequence 1, pings 1, 2, 3, 4, & 5
0106	sequence 1, pings 1, 2, 3, 4, 5, & 6
0201	sequence 2, ping 1
0202	sequence 2, pings 1 & 2
.	.
.	.
.	.
2901	sequence 29, ping 1
2302	sequence 29, pings 1 & 2
etc.	etc.

You will view a total of 72 sequences, each containing 6 consecutive pings. Since it takes 25 seconds per slide, the total viewing time will be 3 hours. However, you will have 3 five-minute breaks, one every 45 minutes.

Contact Reporting Procedure

To permit you to report contacts, you will have a *contact reporting push button* which, when you push it, will record the ping number on which you reported. In addition, you will have a number of *contact reporting sheets* to permit you to record the approximate range and bearing of each contact you report with the push button. The detailed reporting procedure is as follows.

For each sequence, carefully search the display for possible contacts as the slides are presented to you. If, for example, something catches your attention on the 2nd ping of the sequence which you feel you would report as a "possible sub" contact on your ship, *push your contact reporting button* and then immediately write in the first line of the "REPORTS" block of the *contact reporting sheet* the following information:

1. In the first column, labeled "SEQ/PING NUMBER", write the number of the slide (seen in the lower right corner) *during which you pushed the button.*

2. Place a check in the FM or CW columns or both to indicate whether you see the contact in the left 12 beams (FM) or right 12 beams (CW) or both.

3. Estimate which of the 12 beams the contact appears to be in and write this beam number in the column labeled "APPROX BEAM NO.". The beam numbers on your viewing screen will aid you.

4. Estimate the range of the contact from the start of the 20 Kyd A-zone in Kyds and write this range in the column labeled "APPROX RANGE Kyds". The range numbers on your viewing screen will aid you.

5. If you can determine whether the contact has an opening range rate, no range rate, or closing range rate *at the time you report*, place an appropriate check in one of the last three columns of the reporting block. Naturally, to do this you will have to be viewing at least 2 contact echoes at the time you report. If you are not, make no entry in these columns.

To continue the example, you might feel you would *not* report the "something" which caught your attention on ping 2 without further evidence. In that case, simply wait for additional pings to come into view. *Report on the ping number which you feel displays enough evidence that you would report that contact as "possible sub" to the bridge of your ship during routine non-ASW operations at sea.* Do not report *earlier* in the sequence than this, and do not report *later* in the sequence than this. If on ping 6 of a sequence you still do not see enough evidence of a contact, do not report at all.

Once you have reported a given contact in a sequence, you should not report it again in that sequence. However, it is possible that after making a contact report in a given sequence you will see *another* contact, in the same sequence, which you feel would be reported also. In that case, push your contact reporting button *again* and immediately record the sequence/ping number and the range, bearing, etc., information for the 2nd contact on your *contact reporting sheet*, on the line below that used for reporting the 1st contact. Every time you push the contact button, you should make an entry on a line of the contact reporting sheet.

If you have made a report during any sequence, place an appropriate check in Block II of your contact reporting sheet, labeled "CONTACT QUALITY, 1st Target," during the *last* ping of that sequence, when you can see all 6 pings. This will permit you to say whether the reported contact is a "good," "fair," or "marginal" contact after seeing all six pings; or perhaps you would feel the reported contact turned out to be "not a contact," but noise or reverberation marks, in which case you can check that block. *If you have made 2 reports during a sequence,* be sure to place a check in Block III of your contact reporting sheet also, labeled "CONTACT QUALITY, 2nd Target." If you have made more than 2 reports during a given sequence, you should comment on the 3rd and subsequent reports in Block IV of your contact reporting sheet, labeled "COMMENTS." Any additional information on reports 1 and 2 may also be placed in this block if you desire.

If you have made any contact reports during a given sequence, you should have entries in Blocks, I, II, and possibly III and IV on your reporting sheet by the end of that sequence. *At the beginning of the next sequence, turn to the next contact reporting sheet.* Each contact reporting sheet which is used should contain information concerning one and only ping sequence. You do not have to use a contact reporting sheet unless you have made one or more contact reports during a sequence, and it is very unlikely that you will see reportable contacts in all the sequences. But, when you *do* make a report and use a contact reporting sheet, be sure to turn to a new one for the next sequence in which you wish to report.

Because we want *your own* judgment concerning contact reporting, we ask you to refrain from discussions with other sonarmen participating today, until you all have completed the task. Your cooperation in this task is greatly appreciated. If you have any questions concerning these procedures, please feel free to ask the representative in charge of the apparatus.

APPENDIX E

SONAR CONTACT REPORTING SHEET
FOR
THE "LOW" COMMAND ATTENTION CONDITION

I. REPORTS--Make entry for every reported contact of a sequence.							
SEQ/PING NUMBER	FM	CW	BEAM NO.	APPROX. RANGE Kyds	OPENING RANGE RATE	NO RANGE RATE	CLOSING RANGE RATE

II. CONTACT QUALITY, 1st Target			
Good Contact	Fair Contact	Marginal Contact	Not a Contact

III. CONTACT QUALITY, 2nd Target			
Good Contact	Fair Contact	Marginal Contact	Not a Contact

IV. COMMENTS: ______________________________

APPENDIX F

SONAR CONTACT REPORTING SHEET
FOR
THE "HIGH" COMMAND ATTENTION CONDITION

SECRET NOFORM
(when identifier block completed)

I. REPORTS--Make entry for every reported contact of a sequence.

SEQ/PING NUMBER	FN	CH	BEAM NO.	APPROX. RANGE Kyds	OPENING RANGE RATE	NO RANGE RATE	CLOSING RANGE RATE

II. CONTACT QUALITY, 1st Target

Good Contact	Fair Contact	Marginal Contact	Not a Contact

III. CONTACT QUALITY, 2nd Targe

Good Contact	Fair Contact	Marginal Contact	Not Conta

IV. COMMENTS: ______________________________

____

____

IDENTIFIER BLOCK
Do not write in this space

Contact Code ____XXX____ (1) p(D) ____XXX____ (4
Threat Code ____XXX____ (2) p(FA) ____XXX____ (5
Type ____XXX____ (3) Z-Code ____XXX____ (6

SECRET NOFORM
(when identifier block completed)

APPENDIX G

RANKING INSTRUCTIONS AND QUESTIONNAIRE ADMINISTERED TO THE SONAR OPERATORS WHO PARTICIPATED IN THE DETECTION EXPERIMENT

SCENARIO RANKING & QUESTIONNAIRE

Please write your NAME: ________________________________

On the *instruction sheet* (which is the next page), we ask you to *rank order* a set of scenarios (printed on cards). To do this, you can spread the cards out in front of you and arrange them in a single column--but you will probably find it more convenient to actually form three shorter columns side-by-side, so you can easily see all of the cards. When you are satisfied with a particular ordering of the cards, please write the card numbers in the boxes on the reverse side of the *instruction sheet*. Then please answer the questions on the front and back of the last page, which is the *questionnaire*.

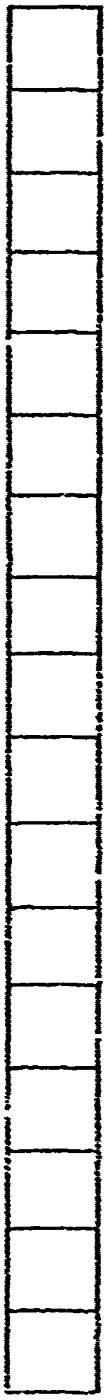
INSTRUCTION SHEET

CONSEQUENCES OF FALSE CONTACTS

For the purposes of this part, please imagine that you are the sonar operator and that you have reported a contact to the bridge as "possible submarine." At least part of the time, a contact reported as "possible sub" turns out to be non-submarine. This situation constitutes a false contact, and can have a variety of consequences. Fuel and/or weapons may be expended, the ship may leave a position in a screen unguarded, the ship's captain may have to be awakened, and so on. The exact consequences of a false contact may be different for different people, and may depend upon the particular operational situation. *Please imagine that you have actually reported a false contact.* Carefully consider what the consequences of reporting a false contact might be *as you see them* for each of the operational situations described on the cards, and place the cards in rank order in front of you so that the situation whose *FALSE CONTACT CONSEQUENCES WOULD BE MOST SEVERE, MOST UNDESIRABLE,* and/or *MOST UNPLEASANT* is at the top, downward situation-by-situation to that whose *FALSE CONTACT CONSEQUENCES WOULD BE LEAST SEVERE, LEAST UNDESIRABLE,* and/or *LEAST UNPLEASANT.* After you have done this, please write the card numbers in the boxes on the other side of this page in the order you have placed the cards.

CONSEQUENCES OF FALSE CONTACTS

CONSEQUENCES
MOST SEVERE



CONSEQUENCES
LEAST SEVERE

QUESTIONNAIRE

FOR EACH QUESTION, CHECK THE ONE RESPONSE THAT MOST ACCURATELY REFLECTS YOUR OWN OPINION.

1. During *present-day non-ASW peacetime operations* (for example, peacetime transiting, non-ASW refresher training, and so on), which do you feel sonar operators *should* do?

____ a. Report fewer false contacts than they do now even though it means increased delay in reporting.

____ b. Continue present reporting practices.

____ c. Report contacts more quickly than they do now even though it means more false contacts.

2. During *present-day ASW exercise situations*, which do you feel sonar operators *should* do?

____ a. Report fewer false contacts than they do now even though it means increased delay in reporting.

____ b. Continue present reporting practices.

____ c. Report contacts more quickly than they do now even though it means more false contacts.

3. Which have you found to be true?

____ a. Many more false contacts are reported during ASW exercises than during non-ASW peacetime operations.

____ b. A few more false contacts are reported during ASW exercises than during non-ASW peacetime operations.

____ c. About the same number of false contacts are reported during ASW exercises and non-ASW peacetime operations.

____ d. Fewer false contacts are reported during ASW exercises than during non-ASW peacetime operations.

4. How much of a problem do you think false contacts would be in the event of a war involving ASW?

____ a. Very serious problem.

____ b. Serious problem.

____ c. Moderate problem.

____ d. Insignificant problem.

5. At the present time, how well do you think the bridge can judge the degree of certainty which sonar has about an initial contact?

____ a. Very accurately.

____ b. Accurately.

____ c. Not very accurately.

____ d. Not at all.

6. How useful do you think knowledge of sonar's confidence or certainty concerning an initial contact would be in decision making on the bridge during a war involving ASW?

____ a. Very useful.

____ b. Moderately useful.

____ c. Of little use.

____ d. Of no use.

7. Does the A-scan presentation of the particular model AN/SQS-26 [AX(R), BX or CX] aboard your ship differ significantly from that of the AN/SQS-26 CX which you have seen pictures of today? In what way?