AD/A-002 307

RESEARCH ON FACTORS INFLUENCING THE INTERPRETATION OF SONAR SIGNALS

Robert R. Mackie

Human Factors Research, Incorporated

Prepared for:

Office of Naval Research

June 1974

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report 776-6	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD/A-002307
4. TITLE (and Subtitle) RESEARCH ON FACTORS INFLUENCING THE INTERPRETATION OF SONAR SIGNALS		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report 1967-1974 6. PERFORMING ORG. REPORT NUMBER 776-6
7. AUTHOR(s) Robert R. Mackie		8. CONTRACT OR GRANT NUMBER(s) N00014-67-C-0537
9. PERFORMING ORGANIZATION NAME AND ADDRESS Human Factors Research, Incorporated 6780 Cortona Drive Goleta, California 93017		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 150-310
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, Virginia 22217		12. REPORT DATE June 1974 13. NUMBER OF PAGES 28
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

18. SUPPLEMENTARY NOTES

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

Sonar Target Classification	Auditory Memory
Sonar Displays	Perceptual Error
Concept Formation	Computer Aided/Managed Instruction
Auditory Perception	Sonar Operators

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

A five-year program of research is described that was concerned with a number of variables that affect the ability of sonar operators to classify underwater sound signals. Five technical reports were issued. The first was an historical review of target clue recognition by operators as a function of various sonar display techniques. The second reported the results of extensive testing of the classification performance of submarine sonar

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. personnel. Data were presented as a function of different target types and different combinations of audio and video displays. The third report was concerned with the use of an auditory memory device, developed under the contract, designed to aid the long-term memory of sonar operators for recognizing various classes of targets. The fourth report addressed the problem of perceptual errors in classifying sonar signals as a function of operator expectations about the nature of the target. The fifth technical report described the use of a small computer, in conjunction with an audio tape library, to provide both basic and advanced training in sonar signal recognition skills. The results of two experimental studies using this Small Computer Auditory Training System are also described.

RESEARCH ON FACTORS INFLUENCING THE INTERPRETATION OF SONAR SIGNALS

Final Report

Robert R. Mackie

Prepared for

Personnel and Training Research Programs
Office of Naval Research
Department of the Navy
Arlington, Virginia 22217

under

Contract No. N00014-67-C-0537
NR No. 150-310

Prepared by

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

June 1974

TABLE OF CONTENTS

	Page
INTRODUCTION	1
CLUE RECOGNITION AND DISPLAY TECHNIQUES	2
CLASSIFICATION PERFORMANCE OF SUBMARINE SONAR OPERATORS	4
MEMORY AIDING AND TARGET CLASSIFICATION	6
STIMULUS ERROR IN THE PERCEPTION OF TARGET CLASSIFICATION CLUES	7
ATTEMPTS TO DEFINE PERCEPTUAL DIMENSIONS OF SONAR SIGNALS	10
USE OF A SMALL COMPUTER IN THE ACQUISITION OF AUDITORY CONCEPTS AND THE DIAGNOSIS OF CONCEPTUAL ERRORS	15
REFERENCES	23

RESEARCH ON FACTORS INFLUENCING THE INTERPRETATION OF SONAR SIGNALS

Final Report

(Contract N00014-67-C-0537)
(NR 150-310)

Introduction

> *The primary purpose of every indicator or recorder used in sonar is, in some way or other, to make the signal perceptible to the observer... the operation of any machine is a joint enterprise in which the instrument is one partner and the observer the other.*
>
> *It is essential that the sonar designer be informed as to the characteristics of observers as well as to the characteristics of the transmitting medium. Sonar, then, is bounded on one side by oceanography, and on the other by psychophysiology....* (Horton, *The Fundamentals of Sonar*, 1959.)

The purpose of the research summarized in this final report has been to study the performance of sonar operator personnel in their most difficult of tasks, signal interpretation or target classification, as it is affected by a variety of display variables and operational procedures. The sonar signals produced by underwater targets, whether the result of echo ranging (active systems), or the mechanical activity of the target itself (passive systems), are among the most complex stimuli that man has been asked to detect and interpret. In the early days of sonar systems, the ear was the primary analyzer of these signals. With the advent of more sophisticated signal processing and display techniques, the eye has also assumed a highly important role in this process. In fact, in some of today's systems, visual pattern recognition is the sole means of analysis; in others, aural and visual analyses play complementary roles.

It was our objective to study the performance of sonar personnel, primarily those aboard submarines, in reaction to both audio and video presentations of a wide variety of representative target signals. In this process, and in accordance with Horton's suggestion, we tried to identify some of the variables associated with both the man and the machine that affect the accuracy of sonar target classification.

Clue Recognition and Display Techniques

The first study conducted under this contract was an historical review of research on the perceptibility of target classification clues as a function of a variety of display techniques that had been developed for active sonar systems. In this review, we noted that the task of classification by aural analysis had become increasingly difficult and that the visual displays then in use provided very limited interpretive clues because they had been designed primarily as detection, rather than classification, displays. A number of display techniques were described that had been shown to be useful for classification, but had not found their way into the design of sonar systems that were then operational. The focus of design attention in active sonar systems had been strongly on improved detection performance, but it seemed to us that the failure to attend to the classification aspect of the problem might well more than offset any gain in detection performance by creating a high false alarm rate. This prediction was certainly borne out by the performance of several systems that were then becoming operational. It is heartening to note that several sonar systems that are now finding their way into the fleet, seven years after this report was issued, reflect considerably more attention to the display requirements for target classification than was true at the time this study began. More progress appears to have been made in this respect with passive sonar systems than with active ones, although sensitivity to the problem is reflected in the designs of the most recent systems of both types.

In our first technical report, we also commented on the evident interest in automating the target classification process and reducing the operator's role to that of a simple monitor. We pointed out that some costly attempts to do this in active sonar systems had proved most disappointing because of the very difficult problem of modeling the adaptive aural and visual processes of man. This remains the situation today in active systems although there has been some success in automating passive sonar classification. But there have been disappointments with passive systems also, largely because of the temptation to *replace* man with the computer rather than use the computer to augment his unique perceptual skills.

In our first report, we pointed out that "man's memory and evaluative capabilities for complex events are admittedly fallible; they must be aided by intelligently designed, special purpose devices and displays. But, somehow, advantage must be taken of man's unique perceptual flexibility to avoid undue reliance on the necessarily rigid criteria of machines for the extraction of meaningful signal characteristics. The consequences of failure to do this will be missed targets, unacceptable false alarm rates, and an increasing number of 'witch-hunts'." It is encouraging that a number of truly interactive systems have recently been built which reflect a much more sophisticated viewpoint toward the solution of this problem than that which prevailed when this project began.

This early review of sonar displays and target clue recognition was reported by Mackie and Parker in Technical Report 776-1: "Active sonar classification and clue recognition: Operator performance and display enhancement techniques (U)," Human Factors Research, Inc., December 1967 (CONFIDENTIAL).

Classification Performance of Submarine Sonar Operators

Our second study represented the first comprehensive attempt to measure the target classification performance of submarine sonar personnel, to measure their accuracy in recognizing clues presumed to be significant for classification, and to analyze the effects of audio and video displays on their performance, both singly and in combination.

Among the important results of this study was the identification of very extensive individual differences among submarine sonar personnel in ability to classify. These differences were associated with the amount of their operational experience, but by no means perfectly, and suggested that considerable improvement in the average level of fleet performance was possible through more extensive training or more frequent refresher training.

Certain techniques for modifying the presentation of the auditory stimulus were shown to enhance clue recognition, although greater accuracy in clue recognition did not always translate into significantly higher classification performance scores. It was shown, however, that performance was better when both audio and video presentations of the target signals were used than when audio alone was used. The fact that improved clue recognition did not always result in more accurate target classification suggested that there might be deficiencies in the process of relating observed clues to a logical classification conclusion. This hypothesis subsequently was tested and verified by the Navy Personnel Research and Development Center (Abrams and Winchell, 1972) who showed that the use of a decision aid in the form of a logic tree significantly improved operator classification performance.

The video display used at the time of this study had some important limitations in its presentation of target clues. This was a systems problem attributable, in part, to certain limitations in signal processing. In the time since this study was reported, significant improvements have been made in the sensor as well as in signal processing and display techniques. The probable consequence is greatly improved clue recognition and classification performance although, to date, we know of no comprehensive study to determine objectively the extent to which this may have occurred.

This study also provided evidence concerning those target classes that were relatively easy, or difficult, for fleet operators to classify. Considerable confusion was shown to occur in the recognition of threat and non-threat classes of targets. The result of these confusions could readily be a high false alarm rate or, at the least, time consuming distractions that could interfere with mission success. (For an excellent review of recent fleet experience reflecting this problem, see Becken, 1974.) To some extent, this problem reflected deficiencies in training. However, a more fundamental problem was the overlapping signal space between targets that are threatening and those that are not. The extent of this problem clearly cannot be assessed except on the basis of a very comprehensive target data base. A sufficient data base did not exist at the time this study was done and we suspect it does not today.

As a result of this study, we recommended that the classification performance skills of all submarine sonarmen, supervisors, and instructors should be determined by objective performance tests at least once a year. We further suggested that advancement in rate should be made contingent in part on satisfactory classification performance test scores. To our knowledge, neither of these recommendations has been implemented although the need seems greater than ever. Recommendations that have been implemented were that classification training be modified to emphasize the correlation of audio and video clues and that target classification procedure incorporate the use of the video display which, at the time of this study, was seldom used.

This research was reported by Mackie, Parker, and Dods in Technical Report 776-2: "Classification performance of submarine sonar operators (U)," Human Factors Research, Inc., September 1968 (SECRET).

Memory Aiding and Target Classification

The classification of sonar signals is often dependent upon memory for complex patterns of sound that the typical operator may not have experienced for months, or even years. The sonar technician typically receives elementary training in target classification during basic operator training and may or may not receive systematic reinforcement of the related perceptual and judgmental skills after he goes to sea. The amount of reinforcement differs widely, depending upon the class of submarine to which the sonarman is assigned and its mission. However, even sonar personnel who frequently analyze sonar contacts may suffer from lack of feedback concerning the true nature of the target since many targets are never positively identified. This can lead, of course, to the reinforcement of perceptual or judgmental errors if the operator assumes his classification is correct when in fact it is not.

Because of this lack of systematic reinforcement, and because the tasks of clue recognition and clue correlation are so complex, it seemed to us that a system which could serve to refresh the memory of sonar operators concerning characteristics of signals typically produced by various classes of targets might be beneficial to performance. Consequently, the next study we undertook was to develop and test an experimental auditory memory aid for passive sonar target classification. The memory aid comprised a 20-channel, random access, magnetic tape playback deck loaded with recorded examples of signals from several classes of targets to which an unknown contact might belong. This provided a convenient means of comparing the signal pattern of an unknown target with any of the selected examples contained in the memory bank.

To test the effectiveness of this device, the classification performance of a group of *experienced* sonar technicians was tested, with and without use of the memory aid. It was shown that the performance of sonar personnel with aided memory was significantly better than with unaided memory; that memory aiding was more effective for some target classes than others; and that subjects who performed below average in the unaided condition benefited more from memory aiding than those who were above average. Since we had developed only a relatively crude memory aiding device, with a very limited store of target reference signals, we concluded that a more refined memory aiding system might have even more beneficial effects.

So far as we know, no steps have been taken to incorporate a memory aiding feature into any sonar system presently under development. This may reflect a de-emphasis of the audio display and recognition of the inherently greater memory features of video displays. We are inclined to believe, however, that the concept has general merit regardless of the display channel; further, we believe that there are today operational circumstances under which auditory memory aiding could significantly benefit performance.

The research concerning memory aiding of sonar operators was reported by Dick, Mecherikoff, and Mackie in Technical Report 776-3: "Enhancing passive sonar classification by aiding the auditory memory of sonarmen (U)," Human Factors Research, Inc., June 1970 (CONFIDENTIAL).

Stimulus Error in the Perception of Target Classification Clues

Training sonar personnel to recognize very subtle target clues in a noise-masked signal is a difficult task. Even the most proficient operators have difficulty describing precisely what characteristics of the signal they are reacting to when they "recognize" a particular clue. Certain classes of targets are presumed to display particular clues on the basis of known mechanical features and operating characteristics. Whether or not such clues are displayed is problematical, however, depending upon many physical variables over which the operator has no control.

We were impressed with the possibility that operators might perceive particular clues on the basis of the kind of target they *expected* to encounter, since mission intelligence plays a significant role in the total classification analysis. (For operational evidence of this problem, see Becken, *op. cit.*) We had observed that target classification instructors often appeared to perceive clues in recorded target signals that were difficult for others to discern. Since the instructors knew the nature of the signal source, the possibility existed that they were "hearing" certain clues whose actual presence might be doubted. This is a phenomenon similar to "stimulus error" that has long been the subject of study by experimental psychologists; it was described as early as 1909 by Titchener (Woodworth, 1958). Stimulus error refers to the fact that an observer's report of a perceptual experience can be influenced by his knowledge or beliefs about the object he is observing. His perceptions of an object (stimulus) may be influenced by a former association, by the general context in which it lies, or by unreflective interpretations (English and English, 1958).

Because of the potential significance of such a phenomenon for sonar target classification, a study was performed to test whether the perception of target clues (and therefore classification accuracy) might be influenced by knowledge or suppositions about the nature of the target. Experienced sonar personnel listened to recorded signals from a variety of sonar contacts and reported the clues they perceived. For half of the signals, valid suggestions were made about the nature of the target; and for the other half, invalid but plausible suggestions were made. The operators, of course, had no knowledge concerning the validity of these suggestions.

The results showed that the perceptions of target clues by experienced operators were indeed influenced by the suggestions made about the nature of the target; further, the frequency with which particular clues were reported was a function of certain stereotypes associated with the different target classes. It was also noted that judgments of whether a target was a typical representative of a class seemed to depend upon unknown characteristics of the signal itself, and to be independent of whether or not the target class under consideration by the operators was the correct one.

The findings of this study have clear implications for instructional procedures in sonar target classification as well as for operating practice aboard ship. Since the study was limited to aurally presented sonar signals, it is not known whether visual presentation of the same signals might have reduced the stimulus error. We are inclined to think, however, that this problem is a pervasive one in surveillance systems. Man has unique capabilities to perceive subtle characteristics of signals in noise, a capability that is hard to emulate through computer software. At the same time, he is susceptible to perceptual error as a result of his expectancies, beliefs, or desires concerning the possibility of encountering a target of interest. Therefore, we believe that this demonstration of a fundamental finding from experimental psychology, in the context of the classification task, has considerable operational significance.

We are not aware of the extent to which the results of this study may have influenced current training procedures, but are inclined to feel that the impact has been minimal. Perhaps this again reflects the increased emphasis on video displays in sonar target classification but, as suggested previously, we expect that the problem may well reside in that information channel as well.

The research concerning stimulus error was reported by Dick, Mecherikoff, and Mackie in Technical Report 776-4: "Susceptibility of experienced sonarmen to stimulus error (U)," Human Factors Research, Inc., August 1970 (CONFIDENTIAL).

Attempts To Define Perceptual Dimensions of Sonar Signals

Sonar signals are commonly assumed (by sonar personnel) to have distinctive sounds associated with the physical or behavioral characteristics of the class of targets from which they come--sounds which can be reliably named. Some of the terms describing the sounds relate directly to physical variables, for example, tones. Others are decidedly subjective impressions; for instance, cargo ship signals are felt to be more "laboring," less "galloping," lower in pitch, and less "charging" than are warship signals. Descriptive terms of both types play important roles in both the instructional process and in operating doctrine for passive sonar classification. Such descriptive terms are often used in training, where the assumption is made that the trainee's attention can be directed to certain qualities of the sound by means of these verbal pointers. The evidence indicates that this practice may be deceptive, sometimes even in the case of terms describing a physically simple characteristic (there are adults who, without specific training, do not know how to recognize a "high-pitched tone" in a complex signal). We were interested in how reliably these and other descriptive terms might be ascribed to representative sonar signals and in determining whether a more fundamental set of descriptive dimensions for the auditory experience might be identified. The identification of perceptual dimensions potentially entails the reporting of subjective experiences by the subjects, a procedure usually shunned because of its inherent unreliability, or else the deliberate variation of stimuli in simple ways along well-defined physical dimensions. The latter was impractical in this case because of the complexity inherent in the sonar signal. To circumvent these problems, the method of triadic comparisons appeared on the surface to be quite promising. In this method, three stimuli at a time are made available to the subject for comparison; the subject indicates which two of the three are most similar and which two are least similar. From each subject's responses, a matrix of stimulus similarities is constructed and any one of a number of methods may be used to cluster the stimuli or factor the similarity matrix. Of course, the usual problem of the factor analytic methods remains: At some point it is necessary to reconsider the stimuli themselves and to attempt to characterize or name the dimensions or clusters and relate them to physical variables.

An additional problem with the method of triadic comparisons is that with even a moderate number of stimuli, the number of triads becomes enormous; the demands on the subject become very heavy, and subject motivation, interest, and reliability can suffer. However, the sonar memory aiding device, developed for earlier work under this contract, made it quite convenient to compare different signals so it was decided to proceed despite these recognized difficulties.

Four pilot studies were carried out using the triadic comparison method to test the applicability of this method for identifying perceptual dimensions of passive sonar sounds. Seven stimuli (recordings of sonar contacts) were used, covering as wide a range of target classes as possible. Naive subjects evaluated the 35 triads, and the resulting similarity matrices were clustered by means of a hierarchical clustering scheme. The resulting clusters showed remarkable consistency across subjects; but in order to check the reliability of the procedure, a second set of seven stimuli was selected and administered to the same subjects. This time there was less consistency, the subjects complained that the judgments were much more difficult, and post-experimental questioning seemed to indicate that part of the lack of reliability was due to the fact that the sound on a given tape loop was not homogeneous throughout the entire loop. In an effort to correct this problem, a 15-second "consistent-sounding" segment was taken from each of the 14 tapes previously used, and these were administered in triads to the subjects. In general, there remained a moderate amount of consistency across subjects in the way the stimuli clustered together, but the reliability was not noticeably improved by using short tape loops.

The relationship of the two or three clusters which tended to emerge to target classes or classification clues was tenuous. The problem, perhaps peculiar to an applied study of this sort, stems from the fact that the clusters, factors, or dimensions which may be most readily identified through the analysis may not relate to the categories of practical interest (in this case, target classes); they are factors which are recognized as irrelevant by experienced operators even without the benefit of dimensional analysis. Some of the most outstanding distinctions between passive sonar sounds as heard by naive listeners are based on prepotent characteristics which happen to be largely unrelated to target class, such as signal-to-noise ratio, overall loudness of the recording, and background noise characteristics. Using trained subjects (experienced sonarmen), however, introduces the genuine danger that similarity judgments will be made on the basis of *inferred target class*, rather than on characteristics of the signal. Furthermore, trained sonarmen, in the formalized structure of an experiment or test, will probably tend to revert to the "standard" characteristics ("nature-of-sound" clues) which they are supposed to use, even though in practice their perceptual and decision processes may be quite different.

Another fundamental problem in the use of the triadic method with very complex stimuli such as these is that the subject may attend to different aspects of the same stimulus at different times and, of course, the judgments of similarity and difference will be seriously affected. Subjects sometimes reported such an attention shift even during their evaluation of a single triad. It is also undoubtedly true that the characteristic attended to depends on the particular triad; for example, "hiss" might be an outstanding characteristic of one of the sounds in a particular triad, but "hiss" might be completely ignored in another triad if all (or none) of the sounds happened to have "hiss." If there are many characteristics in each sound to which the subject may attend, it will require a very large number of stimuli and many replications of the triads to identify all the dimensions, particularly if the dimensions of practical importance are not the most obvious ones. However, the number of triads generated by even a rather small number of stimuli confronts the subjects with an enormous, fatiguing task and makes the method prohibitive. This is not to imply that using multidimensional scaling or clustering techniques for the identification of perceptual dimensions in sonar signals is impossible, but only that a successful effort to use triadic comparisons to that end would be massive, and would require a very large number of subjects each working for many hours.

Following this somewhat disappointing experience with triadic comparisons, we decided to try a simpler, more direct approach. A list of approximately 30 descriptive words was compiled by three of the research staff. Some of these words were taken directly from the vocabulary used in the sonar aural analysis course, while others were suggested by staff members after listening to a great variety of sonar tape recordings. Following each descriptive word in the list was a 10-point scale, and the subjects were instructed to indicate the strength of that particular characteristic in the sound to which he was listening. The stimuli used were the fourteen 15-second loops prepared previously; these were used so that a direct comparison could be made with the triadic data. Again, subjects were used who had no sonar experience; part of the objective was to discover what kinds of terms naive subjects were willing to use in describing what they heard. More specifically, we wished to find out to what extent these subjects would use the vocabulary regularly used in sonar training.

The results were highly variable across the subjects. There was little inclination among these subjects to choose the sonar school terminology. Of course, it may be argued that one function of the sonar training is to attach the proper verbal labels to the appropriate aural experiences and that naive subjects cannot be expected to use the vocabulary properly. Although this argument is probably valid, it is also true that these terms are used in training as pointers to a particular characteristic of a complex sound. The instructor, having used the terms for many years, tends to overrate the trainee's ability (as well as his own, sometimes) to associate the terms with exactly the correct characteristic.

There were four terms in our list, however, which seemed to receive consistently high ratings for some stimuli and low ratings for others (Hiss, High-Pitched, Pulsing, Squeal). For the sake of a rough model, these four descriptive words were treated as orthogonal dimensions, the 14 stimuli were considered in pairs, and a Euclidean distance was calculated between the two stimuli in each pair. When the stimuli with small distances between them were considered to be clustered together, there was good agreement between these clusters and the clusters previously obtained through triadic comparisons. Furthermore, by considering all 14 stimuli together, it was possible to construct some triads which had not been previously tested, and to predict on the basis of the "distances" between the stimuli how the similarity and difference judgments would be made within the triad. Although all these data were derived from pilot studies with very few subjects, the consistency encouraged further investigation.

The above study was then repeated using experienced sonarmen, with two changes. The stimuli were selected this time in order to represent each classification category more equally, and the response sheet was simplified. The results essentially paralleled those of the first study: There was some distinctiveness in the profile for each target class considered, but there was too much overlap on each "dimension" (descriptor) for the technique in its present state to have value as a classification aid.

Thus, the experimenters' growing suspicions that perceptual dimensions, in any theoretically or mathematically pure sense, would remain elusive were strengthened; while the modest success of the terminological study gave hope that careful attention to the processes of discrimination, concept acquisition, and verbal labeling in aural classification training might yield a set of "dimensions" which would serve the practical end of improved target classification, and might eventually be related to more fundamental auditory processes.

Because of the exploratory nature of these studies, they were not reported in a separate technical report. However, they are discussed by Mecherikoff in Technical Report 776-5: "Concept learning in the aural analysis of passive sonar signals," Human Factors Research, Inc., 1974 (CONFIDENTIAL).

Use of a Small Computer in the Acquisition of Auditory Concepts and the Diagnosis of Conceptual Errors

The failure for methodological reasons to discover the perceptual dimensions of sonar signals that relate to classification performance led us to try an entirely different approach to the problem. It is a reasonable assumption that these dimensions, or at least their relevancy, are determined largely by training and experience. Therefore, another way of discovering perceptual dimensions might be to manipulate the training conditions assumed to produce them. The problem of sonar classification training was viewed as one of developing in the trainees a set of auditory concepts, some related to specific clues or characteristics of the sound, and others related to the overall quality of the sounds of the various target classes.

Very little of the research on concept acquisition has made use of nonverbal auditory stimuli, and most studies which have used auditory stimuli have questionable application to sonar training since they typically used simplistic stimuli in experimental sessions of very limited duration. We felt that experimentation was needed which would simulate actual training conditions (although in a highly controlled fashion), which would use carefully selected tape recordings of actual sonar signals, and which would extend over a sufficient period of time to enable considerable concept acquisition to take place. To administer the complex experimental conditions and record comprehensive data, a small but sophisticated computer-controlled auditory training system was developed (Small Computer Auditory Training System).

The SCATS consists of a small digital computer (12K of core) and magnetic tape unit that presents the stimuli (sonar signals), administers the training sequences, and records all subject responses to the signals presented. Up to eight operator terminals can be operated simultaneously and independently, each consisting of a keyboard, a limited visual display for feedback and directions, and headphones. The taped audio signals are stored in endless loop cartridges, and up to 32 high fidelity channels are available instantly to any terminal. The audio system has excellent fidelity, the system response being flat from 20-25,000 Hz ± 2 db, and the headphones from 20-15,000 Hz ± 3 db.

Software development included a sophisticated executive program, similar in concept to IBM O/S 360, and a flexible interpretive language designed for writing the training sequences.

In addition to the usual benefits of CAI, the design of this system provided maximum efficiency in training audio concepts, provided an amount and variety of carefully guided listening experiences many times greater than that available with conventional training methods, and assured that the trainees' attention was focused on the proper feature of the complex sonar signal by requiring meaningful responses to each stimulus.

Although the system was not designed as a training device *per se*, it is clear that it could be used to reduce the need for instructor manpower in the laboratory portions of operator training. Its cost is modest since it does not require CRT displays, and does not attempt verbal "conversations" with the trainees.

The basic functions of the system are as follows:

1. To present auditory signals to up to eight operators working simultaneously but independently;

2. To accept responses from the operators via a keyboard terminal and evaluate and act upon these responses;

3. To provide feedback and direction to the operators through a dual display consisting of a three-digit numeric readout and a set of 16 lights indicating relevant messages printed on replaceable masks;

4. To keep a comprehensive log of all stimulus conditions presented to the trainees, and all trainee (and instructor) actions, for future detailed evaluation of trainee performance and of training effectiveness.

Due to the time required for getting this small but sophisticated system to function properly, and due to an unexpected decrease in subject availability at the Fleet ASW School, only two studies of concept acquisition were eventually completed using this system. The first study was directed at Navy personnel who did not have prior target classification experience or training; the second was directed at senior Petty Officers at the Fleet ASW School who were experienced, but in general did not consider themselves experts in classification. In the first instance, we were interested in initial concept acquisition; in the latter, we wished to identify the existence of possible erroneous concepts and to prescribe, and give, corrective training based on individually diagnosed deficiencies in performance (concept modification).

In the experiment on initial concept acquisition, three different procedures were employed:

1. *Target Class Recognition Training.* In this condition, examples of six target classes were presented as the concepts to be learned. This condition afforded the greatest opportunity for the subject to respond to the overall quality of the sound and to react to any subtle patterns that might have been common to many members of the class but which may be difficult or impossible to verbalize. Of the three approaches used, it was the most Gestalt-like.

2. *Clue Training and Class Recognition Training.* This condition most closely approximated the kind of conceptual training currently given in the ASW School. Attention was focused on the recognition of specific target clues, and several clues were presented as concepts to be learned. The clues were related in a general way to associated target classes, although no specific logical procedure for combining them into a specific classification conclusion was provided.

3. *Clue Recognition Training with Classification Logic.* In addition to training on the recognition of specific clues, this condition included instruction on the use of a formal decision logic which systematically took the operator to a particular classification conclusion based on his perception of the clues. This condition was the most analytic and mechanical of the three approaches to classification.

Each of the three experimental groups experienced a total of 15 hours of training and testing using the SCATS. Although only three hours of the training were devoted to clue recognition, the trainees made an average score of 76% correct in determining the presence or absence of eight different clues in representative sonar signals. Some clues were easier to learn than others, the number of correct judgments ranging from 62% to 97%. In general, these percentages were about the same whether the clue was present or absent. It is interesting to note that the clue recognition scores of these inexperienced subjects were as high after this short period of training as those of the experienced sonar Petty Officers tested in the second study (68% to 92%). This evidently testifies to the highly efficient training that SCATS makes possible. About 75% of all the training time was spent actually listening to sonar contacts and making perceptual judgments.

The data concerning performance in actual target classification were less impressive. All three groups made some progress in learning to classify; but the final level of achievement was only about 10% above chance, except for Group 3, which performed 19% above chance. Since Group 3 was the only one directly led through a structured logic process for relating target clues to a decision outcome, this result suggests that the use of logic trees may be one means of improving classification performance. It also confirms the earlier findings of Abrams and Winchell (*op. cit.*), who originally developed the logic tree used in this study, in their tests of experienced sonar operators.

In the second experiment, experienced submarine sonar personnel were first given a diagnostic pretest by SCATS, the results of which were immediately processed to identify the kind of classification errors each operator had made. A program of remedial instruction was then embarked upon, also administered by SCATS, which was individually tailored to the performance weaknesses of each sonarman. Pretest-posttest comparisons showed that even a very short (1-day) refresher training course had an overall beneficial effect upon target discrimination capability. The data revealed substantial individual differences in the amount of time operators used in classifying the training items and showed, in general, that the more items they processed during refresher training, the higher were their posttest scores. This fact alone argues for the use of a system like SCATS to replace rather cumbersome procedures presently in use.

During remedial training, the computer routine required the operators to classify each sonar contact on four separate occasions. Each operator first made a classification based on his immediate impression after a 10-second presentation of the signal. Following this, the same signal was turned on for up to 5 minutes for subjective analysis and reclassification. Generally speaking, these responses were also made quickly, usually in less than one-half minute. Following this, the operator was asked to judge the presence or absence of eight standard clues, one by one, and to get a beat count. With the results of these analyses displayed, he classified the target again, whereupon the correct beat count and target clues were displayed. He then classified the target a fourth time, taking the correct evidence into account if it differed from his own.

It is notable that classification performance on the training items increased very significantly as a function of these successive steps, each of which provided the operator with more information than he had for the preceding steps. The total percentages correct for the four successive classifications were 39%, 44%, 49%, and 57% across all stimuli and all subjects. This showed that, on the average, performance was poorest with immediate recognition, better with an unguided analysis of the target signal, better still when the operator was forced to go through a systematic clue analysis, and best when the correct clues were actually presented. This last result indicates that failures in clue recognition, in addition to logical errors, are responsible for errors in target classification.

There were also some noticeable shifts in the recognition of certain classes of targets, particularly high threat targets, with which these personnel had had little prior experience. For example, in the case of one important target category, classification accuracy improved from 18% to 48%. This clearly showed that the emphasis on this class of target during remedial training had a facilitating effect on concept acquisition of that particular class. Unfortunately, it came at the expense of some loss in accuracy in classifying other types of targets. (Analysis indicated, however, that it was not the simple result of developing a response bias.) Further work is necessary to determine what program changes can be introduced to achieve significant improvement for all target classes. Perhaps it is just a matter of increased time, or better distribution of training time, with the SCATS.

It has already been pointed out that a brief period of clue recognition raining produces clue recognition performance in novices comparable to that of experienced sonarmen. In addition, an analysis of target class confusions for the two groups indicated a substantial similarity between novices and experienced sonarmen in the types of classification errors made. In short, after rather brief training, novices were beginning to behave in significant ways like experienced operators. Results of this kind should whet the appetite for more extensive studies of the very promising analytic and training capability that computer-based systems like SCATS represent. The system, and detailed results of the two concept formation studies, are presented by Mecherikoff in Technical Report 776-5: "Concept learning in the aural analysis of passive sonar signals," Human Factors Research, Inc., 1974 (CONFIDENTIAL).

New sonar systems presently under development will demand increased sophistication, rather than less, in the interpretation of underwater sound signals. The fundamental issues addressed in this study will continue to be important to successful operation of the Navy's sonar systems, though the specific problems may change as a function of new display and signal processing techniques. Happily, there appears to be increasing awareness throughout the Navy of the very complex tasks faced by the operators in the target classification process. But the need for continuing research and evaluation in this problem area was never more urgent.

REFERENCES

Abrams, A. A., & Winchell, J. D. Evaluation of the physical analysis classification training system (PACTS) (U). Naval Personnel and Training Research Laboratory, San Diego, California, 1972 (SECRET-NOFORN).

Becken, B. A. Contributions of ASW to carrier task force defense (U). *Journal of Defense Research*, 1974, *6B, No. 1* (SECRET).

English, H. B., & English, A. C. *A comprehensive dictionary of psychological and psychoanalytical terms.* New York: Longmans, Green & Company, 1958.

Navy Training Research Laboratory. Fundamental audio classification training system. Naval Personnel and Training Research Laboratory, no date (SECRET-NOFORN).

Woodworth, R. S. *Experimental psychology.* New York: Henry Holt & Company, 1938.

TECHNICAL REPORTS PRODUCED

Mackie, R. R., & Parker, E. L. Active sonar classification clue recognition: Operator performance and display enhancement techniques (U). *Human Factors Research, Inc., Technical Report 776-1,* December 1967 (CONFIDENTIAL).

Mackie, R. R., Parker, E. L., & Dods, W. D. Classification performance of submarine sonar operators (U). *Human Factors Research, Inc., Technical Report 776-2,* September 1968 (SECRET).

Dick, R. A., Mecherikoff, M., & Mackie, R. R. Enhancing passive sonar classification by aiding the auditory memory of sonarmen (U). *Human Factors Research, Inc., Technical Report 776-3,* June 1970 (CONFIDENTIAL).

Dick, R. A., Mecherikoff, M., & Mackie, R. R. Susceptibility of experienced sonarmen to stimulus error (U). *Human Factors Research, Inc., Technical Report 776-4,* August 1970 (CONFIDENTIAL).

Mecherikoff, M. Concept learning in the aural analysis of passive sonar signals (U). *Human Factors Research, Inc., Technical Report 776-5,* June 1974 (CONFIDENTIAL).

Mackie, R. R. Research on factors influencing the interpretation of sonar signals. Final Report. *Human Factors Research, Inc., Technical Report 776-6,* June 1974.