UNCLASSIFIED

[redacted] MOST Project -3

Interim Report

LEVEL II

ANALYSIS OF ASW CONTACT CLASSIFICATION PROCEDURES

Prepared for:

Chief of Naval Personnel
Bureau of Naval Personnel (Pers-C133)
Department of the Navy
Washington, D. C. 20370

Contract No. N00022-67-C-0172

UNCLASSIFIED

[redacted]

UNCLASSIFIED
[redacted]

(14) BSD [redacted]-67-439

(9) Interim Report,

(6) ANALYSIS OF ASW CONTACT CLASSIFICATION PROCEDURES

(15) N00022-67-C-0145

Prepared for:

Chief of Naval Personnel
Bureau of Naval Personnel (Pers-C133)
Department of the Navy
Washington, D. C. 20370

(11) 28 Nov[redacted]67

(12) 59p.

ACCESSION for	
NTIS	White Section ☒
DDC	Buff Section ☐
UNANNOUNCED	☐
JUSTIFICATION Per Letter on File	
BY	
DISTRIBUTION/AVAILABILITY CODES	
Dist.	AVAIL. and/or SPECIAL
A	

Prepared by:

Dunlap and Associates, Inc.
One Parkland Drive
Darien, Connecticut 06820

ie

C.,
la-

GROUP 4
Downgraded at 3 year intervals, declassified after 12 years.

118620

671215-0400

[redacted]
UNCLASSIFIED

UNCLASSIFIED

[redacted]

ACKNOWLEDGEMENTS

Many individuals on the Dunlap and Associates staff contributed to this study and report. Michael Nacht assisted significantly in the detailed development of the analysis and was in charge of all of the work done on the computer. Howard Weiss did most of the programming.

Joseph Wohl, in addition to handling all of the supervisory administrative chores, provided constant guidance and penetrating comments which clarified and sharpended many of the concepts and much of the presentation. Stanley Barnes (who prepared his comments on the pentultimate draft while recovering from an illness) made the basic suggestion which led to the proposed procedural changes and the supporting analysis, and contributed many other helpful detailed suggestions. Al Van Acker was constantly helpful in spotting errors, vaguaries, and misconceptions. Thanks are also proffered to all of the Dunlap and Associates study team members who not only sat patiently through long briefings and conferences as the analysis developed, but also provided penetrating and constructive advice.

Finally, a special thanks to Miss Susan Sullivan who coped skillfully and patiently with all of the problems of preparing the various drafts and final copy of all of the material in the report.

UNCLASSIFIED 671215-[illegible]

[redacted]

ABSTRACT

ASW contact classification operations are analyzed as a two-step process. In the first step, a classification "system," which includes the sonar operators, reduces the multivariate sonar input information to one of a simpler set of "observations." In turn, these observations are used by a "decision maker" in selecting the classification category. The systematic transformation of information is described by a "receiver operating characteristics" (ROC). A rational way of using the ROC in decision making is to apply a "maximum decision rule" which ensures that the resulting risk is not greater than the least upper bound of all of the possible risks in the given situation. The value of the system is defined as the difference between the maximum risk obtained when the system is not used and the maximum risk obtained when the system is used.

(C) This measure of system value is used, together with ROC's obtained by the Defense Research Laboratory of the University of Texas, Austin in deriving two results. First, it is shown that the classification activities of the sonar operator can improve ASW effectiveness by an amount larger than the cost of the training and other measures needed to develop the required operator skills. Second, it is shown that ASW effectiveness can be improved by modifying the sonar operators' classification outputs. As a consequence, it is recommended that the planned experimental follow-on study be implemented, that the procedural modifications be investigated in detail, and that the analysis be extended to an investigation of significant questions not yet considered.

TABLE OF CONTENTS

1. SUMMARY AND CONCLUSIONS

1.1 Introduction

In accordance with Contract NO0022-C-0172,* Dunlap and Associates, Inc., has carried out "an analysis that determines the sensitivity of contact classification accuracy to sonar operator inputs. . ." This report describes the analysis, the results, and the consequent recommendations for further action.

It is demonstrated that the sensitivity is high so that the follow-on research already proposed should be initiated. In addition, there is an unanticipated result: it is demonstrated that a relatively simple change in the sonar operator's reporting procedures can produce a significant increase in ASW effectiveness. It is recommended that investigation into this possibility be initiated.

1.2 Summary of the Analysis

The salient points of the analysis are as follows:

(C) 1) The classification of sonar contacts proceeds in two steps. In the first step, a "classification system" converts received sonar signals into a form which a "decision maker" can see or hear. On the basis of these observations and any other available information he may consider relevant, the decision maker then either proceeds as if the contact is a submarine or proceeds as if the contact is not a submarine.

(C) 2) In the current systems, the sonar operator plays a dual role: he functions as part of the system in converting signals into observable form (for example, by identifying cues), and as a decision maker by using his observations in arriving at either the statement that "the contact is a possible sub" or the statement that "the contact is not a sub."

(C) 3) The CO** functions as a decision maker. The sonar operator's output is the observation which he then combines with other information in selecting an appropriate tactical response. Designating the CO as the terminal decision maker is an analytical convenience: if he merely passes a refined classification decision to the screen commander or the OTC he plays a dual role analogous to that of the sonar operator, and the OTC becomes the terminal decision maker. Interposing any number of intermediate steps of this kind does not change the basic analysis or the conclusions.

---

* Item 3
** Commanding Officer

4) *The decision maker's procedure is described as the selection of a "decision rule"; that is, a rule which associates each possible observation with one of the available tactical responses. The decision rule can also be visualized as a way of stating in advance just what will be done if certain system outputs are observed.*

5) For any particular system and any particular decision rule two types of errors can occur: a) erroneously proceeding as if the contact is *not a submarine* when in fact it is and b) erroneously proceeding as if the contact is a submarine when in fact it is not. Numerical measures called "losses" are assigned to indicate the degree of undesirability of these events. Expected (or "average") losses are called "risks."

6) The probability that each of these errors will occur depends upon the characteristics of the system and the decision rule used. In general, when the decision rule is changed, one of the errors increases and the other decreases. The relationship between these two types of errors defined over all of the possible decision rules is called the "receiver operating characteristic" or "ROC." Since it covers all possible rules, the ROC is characteristic of the system. It summarizes all of the available information pertinent to the evaluation of system performance.

7) Risks are calculated for a given ROC (*that is, for a given system*) and each possible decision rule. There is one decision rule, called the "maximin decision rule" with this very important property: when the maximin rule is used, regardless of any other aspects of the circumstances, the risk will not exceed a particular finite value called the "maximin risk"; when any other decision rule is used, the limiting value of the risk may be larger than the maximin value, and will certainly not be smaller. The maximin rule is "optimal" in that it puts the lowest ceiling over the possible risks.

8) There is an analogous maximin risk obtainable when the classification system is not used; that is, when the classification is decided on the basis of all available information other than that derived from the classification system. The difference between the maximin risks with and without the system is defined as the "general value of the system."

(C) 9) System values were calculated for the case in which the assignment is in a screen protecting a merchant ship convoy. This is a limiting situation in that the losses will usually be higher than those assumed both for this assignment and for all other possible ASW assignments. For this situation, and available ROCs for systems incorporating either unskilled or highly skilled operators, the calculated increase in general value indicates that the cost of training which might be required to achieve this increase would be less than the replacement value of escorted ships saved by use of a skilled rather than an unskilled, operator in the course of 1000 miles of convoy duty.

(C) 10) The value of a system in which the sonar operator's outputs are either "possible sub" or "non-sub" is compared to the value of a system in which the output is the operator's subjective estimate of the likelihood that the contact is a sub or a non-sub. The latter system is usually more effective than the former, and never less effective.

1.3 Conclusions

(C) 1) It has been demonstrated that the "target classification accuracy" achievable in ASW operations, judged in terms of the consequent selection of appropriate tactical measures, is significantly sensitive to the "sonar operator inputs." Further, the sensitivity is high enough to ensure that costs of reasonable measures to improve operator effectiveness (for example, training, personnel selection, improved procedures and equipment) will more than be recovered as a result of the improved general effectiveness.

(C) 2) A significant improvement in overall effectiveness can be achieved by changing the sonar operator's output from the present choice between "possible sub" and "non-sub" to a choice from a larger number of alternatives (perhaps ten) indicating his feeling of certainty about the classification. If this change is introduced, it will also be necessary to develop new procedures for the CO. If the final selection of an action is the responsibility of the OTC* rather than the captain, the number of output alternatives available to the captain should be increased similarly. The reasoning applies up the heirarchy to any terminal decision maker.

1.4 Recommendations

1) The proposed experimental investigation into the relevant aspects of clue ambiguity should be implemented: this is the first step in capitalizing on the conclusions of this study.

2) Concurrent with the experimental investigations, the applications of this analysis should be expanded to explore all classification factors which may have a significant effect on ASW combat effectiveness.

(C) 3) The development of practical procedures for modifying the sonar operator's output from two alternatives to many alternatives and the CO's utilization of that output should be initiated.

---

* or screen commander

2. QUALITATIVE PERSPECTIVE

2.1 Historical Background

(C) The study described in this report originated in an earlier preliminary investigation into possible procedures for measuring sonar operators' performance and the possible use of such measures in evaluation, selection, and training procedures designed to improve the performance.[1] "Classification" was one of the operator functions examined. In performing this function, an operator or team of operators observes the outputs of three devices. The audio signal is presented directly on a head set, a visible representation of the signal is presented on a cathode ray tube, and a more permanent representation of successive signals is traced out on a pen recorder. On the basis of these observations, the sonar operator decides that the contact is either a "possible submarine" or a "non-submarine" (with "submarine" usually shortened to "sub"). This process was analyzed by application of "decision theory," an established branch of applied mathematics. In this analysis it was demonstrated that an operator's effectiveness in performing this function is limited by two factors:

- the ambiguity of the observations
- the significance of the operator's decision in the selection of the appropriate tactical measures.

In an ideal system, the observations are not ambiguous: certain observations occur only when the contact is a submarine and the rest of the possible observations occur only when the contact is a non-submarine. In actual systems, some observations are ambiguous; that is, they occur either when the contact is a sub or when it is a non-sub. Such ambiguity clearly limits the potential performance of an operator since he cannot achieve more accurate discriminations than those inherent in his basic information. In measuring and evaluating classification performance, therefore, it is necessary to distinguish between inability to discriminate resulting from poor operator performance (for example, failure to observe accurately)* and the inherent limitations introduced by the signals, instruments, and procedures.

The second limitation which we call "impact" is apparent as soon as it is recognized that isolated information cannot be evaluated: the significance or value of information must be judged in terms of its effect on explicit, observable actions. In particular, tactical information such as the classification of a sonar contact must be evaluated in terms of the effect of this information on the choice of specific tactical actions such as attacking,

---

* See, for example, the work done by the U. S. Naval Personnel Research Activity, San Diego, in clue extraction capability of sonar operators[13, 14, 15, 16]

alerting other ASW units, or evading. (When the contact is not a submarine, ignoring it is usually the most appropriate "action.") The sonar operator does not select the tactical action: this is the responsibility of the captain or in more complex situations of a force commander (for example the OTC). When the captain makes this selection on the basis of factors other than the sonarman's classification or on the basis of classification information other than that provided by the sonar operator, the sonar operator's classification has less value. In the extreme case, the classification information from the sonar operator is completely ignored, and it then does not matter if this information is correct or incorrect.

(C) In the original study, we concluded that reasonably complete quantitative knowledge of these two limiting factors was a prerequisite to any effort designed to measure classification performance or to increase its effectiveness. Accordingly, it was recommended that additional study be undertaken starting with an analysis of the impact of the operator's classification activity on the effectiveness of the selection of tactical measures. Stating the same concept somewhat differently, we suggested investigating the "sensitivity" of total system performance to operator classification performance. Are the constraints set by the two limiting factors so tight that the total system performance is insensitive to the performance of the classification subsystem? (If so, it is reasonable to concentrate on operator functions such as detection, for which the overall performance is highly sensitive to variations in subsystem performance.) It was further recommended that detailed study of the ambiguity phenomena be undertaken if the sensitivity analysis showed that the sonar operator's classification performance has a significant effect on the overall system performance. These proposals were accepted, and this report describes the results of the initial sensitivity analysis.

(C) Before discussing the study proper, it is important to describe one other historical factor. Although adequate quantitative information about the limitations was not, and is not, available, others have also been investigating aspects of these problems in much the same way. In particular, the Defense Research Laboratory (DRL)* has been pioneering in illuminating the first of the limitations, ambiguity. (Specific references are cited subsequently.) Their work centers about the use of a relationship known as the "Receiver Operating Characteristic" or ROC. As a consequence of ambiguity, two kinds of errors can occur: a non-sub can be erroneously classified as a sub, and a sub can be erroneously classified as a non-sub. The first type of error is usually referred to as a "false alarm." In any particular system the frequency of occurrence of one type of error is decreased when the frequency of the other type of error is allowed to increase. As an extreme example, suppose every contact is

* Of the University of Texas, Austin, Texas

classified as a sub. Then all non-subs will be classified as subs, and the probability of a false alarm will be unity. At the other extreme, if no contact is classified as a sub, no non-subs will be classified as subs and the false alarms will never occur, while the second type of misclassification will certainly occur.

In radar detection systems (the first to be described using this nomenclature) the probability of a false alarm is completely adjustable by turning a knob (setting threshold) and accordingly of no direct interest to the system designer. Rather, the relationship between the two types of errors for all possible false action settings (that is, the ROC) is the significant design parameter.

(C) The early results achieved at DRL have demonstrated that very useful results can be achieved by applying these concepts to ASW classification systems, where the "system" includes the equipment, the operators, and any procedures used by the operators in handling the equipment and interpreting the observations. In ASW systems, the false alarm probability (probability of misclassifying a non-sub as a sub) can be adjusted to accommodate different operational circumstances, though the adjustment is not achieved simply by turning a knob. Accordingly the system designer is interested in the ROC which tells him things like "System A is better than, or at least as good as, System B for all false alarm probabilities." More specifically it tells him "if System A and System B are adjusted to have equal false alarm probabilities, the probability of the second type of error is lower for System B than for System A." DRL has developed and is refining a procedure for measuring the ROC of any ASW classification system. This is obviously very useful to the system designers. Its importance in our study is discussed in the next section.

2.2 The Significance of the ROC Concept

(C) The fact that an ROC can be developed for any given ASW classification system provides a conceptual basis for clearly distinguishing between the two types of system limitations discussed in the preceding section. All of the phenomena associated with ambiguity are comprehensively and compactly represented by the ROC. This includes insensitivity of sensors to signal characteristics which differ for different types of contacts, loss of significant data resulting from deliberate or unavoidable signal filtering, and operator errors resulting from misreading observations, misinterpreting observations, or just wooly thinking. Thus, all of the factors resulting in misclassifications can be investigated by determining the associated ROCs. In particular, DRL has already demonstrated that this technique will distinguish clearly between systems using operators known to be "naive," "average," or "expert." The effect of training, or of altering displays, or of any other operator-connected factor can also be investigated in the same way.

This advances our study in two ways. First, we can conduct sensitivity analyses without worrying initially about the working details of specific systems: all of these details are conveniently summarized in the ROCs. In other words, we can proceed as if the ROC is the system since it tells us all we have to know about ambiguity when investigating sensitivity. It follows that our sensitivity analysis is concerned primarily with impact. Second, the concept of the ROC provides the specific mechanism to be used in the proposed subsequent study of ambiguity. Further, it makes it possible to start investigations of the key human phenomena immediately since the validity and utility of the experimental techniques have already been established.

2.3 The Role of the Ideal Decision Maker

In carrying out our sensitivity analysis, we continue to use the concepts and techniques of decision theory. In taking this approach we recognize that this formulation does not provide a realistic description of the way in which real decisions are made in real tactical situations. The deficiencies of decision theory as a descriptive structure have been frequently pointed out. For one thing, direct use of the theory requires knowledge and understanding of several probability distributions. Although human beings are capable of some sort of subjective weighing of chances (or odds) in uncertain situations, there is no evidence that these subjective estimates bear any real relationship to the probability distributions of interest. In addition, there is one particularly troublesome distribution, the so-called "a priori distribution." There are valid grounds for maintaining that this information is never (or almost never) available in useful form in real situations. Another point is that direct use of the theoretical concepts requires other information not generally available in real situations; namely, knowledge of the relative desirability of the consequences attending various possible situations. As one example, what are the relative consequences, expressed in some quantitative terms, of each of the two types of errors discussed previously; that is, is one twice as bad as the other? three times? etc. Finally, it is probably true that even if all of the required information were available, real people--as distinguished, perhaps, from mathematicians--would never use it in the logical format employed in the theory.

In the preceding paragraph we have deliberately painted some aspects of decision theory in the worst possible light because it is our contention that these considerations do not affect the validity of the analysis as we have used it. In our approach we accept the artificiality of the theory: in fact, to simplify discussion, we introduce the concept of an "ideal decision maker." This is an hypothetical individual who is, first, completely informed about all of the relevant probabilities and utilities, and, second, completely rational in that he uses the information in the ways prescribed by the theory. As a consequence, he makes the best possible use of the system being studied. Except for occasional good luck, a real decision maker will never do better and will usually not do as well. Therefore, our ideal decision maker represents the goal for which humans can strive. Further, when we evaluate the system (or the information produced by the system) using the theory, we recognize that this value is the upper limit to that which can actually be obtained.

Our ideal decision maker can also be looked at from a slightly different point of view. He is the "standard" decision maker who can be inserted into different systems to be compared. In this role, he may do better than the average real person, but he introduces no bias and consequently provides a mechanism for obtaining valid comparisons between different systems.

(C) Finally, there is a practical utility in using the theory which is demonstrated after the fact. The theory identifies certain procedures as "optimal" and it is demonstrated that these procedures are best regardless of the knowledge (or ignorance) of the real decision maker. There are practical ways of recognizing and implementing these procedures. The result is that the use of the theory has led to some practical recommendations of considerable importance despite its limitations as a descriptive construct. These recommendations and the demonstration of their validity were not a result which was anticipated when the sensitivity analysis was originally proposed.

2.4 Decision Rules

The sensitivity analysis centers about the identification and selection of "decision rules." A simple example--deliberately not taken from the ASW situation to avoid the possibility of being confused by preconceptions--provides a useful introduction to the relevant concepts.

Suppose a vehicle with a legal right to ignore traffic lights (say a fire engine) is approaching an intersection at which there is a traffic signal normally tripped by approaching cars. The driver will observe that the light is either red, yellow, or green. As he approaches the intersection, the driver will have to decide whether he should slow down in preparation for a stop or to maintain speed. We call the first action "stop" and the second "go."

CONFIDENTIAL

The driver's problem, of course, is that if there is cross traffic which does not yield the right of way and he does not stop, he will be in an accident which is undesirable in its own right, and doubly undesirable because it makes it impossible to complete the mission (fighting the fire). On the other hand, if he stops unnecessarily he loses valuable time.

In decision theory, the process is structured as follows. The driver makes an OBSERVATION: the possible observations are that the traffic signal is red, yellow, or green. The driver's decision consists of selecting an ACTION: the two possibilities are stop and go. There are two possible STATES: "yes" there is cross traffic moving into the intersection and "no" there is no such traffic. The OBSERVATIONS, ACTIONS, and STATES are the basic elements of the decision process.

The actual decision process is described as the selection of a DECISION RULE. Here are some examples:

1. STOP if YELLOW, RED, or GREEN
2. STOP if YELLOW or RED; GO if GREEN
3. STOP if YELLOW; GO if RED or GREEN
4. GO if YELLOW, RED, or GREEN

It is some help that the color of the traffic signal is an indication (somewhat ambiguous) of the STATE. If GREEN, there is no cross traffic (provided that the signal is working properly). If RED, the signal was triggered by cross traffic, but it may have already cleared the intersection. If YELLOW, the signal was triggered by cross traffic preparing to move into the intersection. These relationships help evaluate the decision rules: No. 1 is unduly conservative, wasting time needlessly and No. 4 is unduly rash, risking an accident unwisely. The choice between Nos. 2 and 3 is not clear: No. 2 may lead to a needless loss of time and No. 3 may lead to an accident. A rational choice between the two involves an estimate of the likelihood that the undesirable event will happen (no traffic on RED when No. 2 is used or cross traffic on RED when No. 3 is used) and the seriousness of the consequences (lost time, accident) in each instance. The combination of likelihood and consequence is called "risk." The rational procedure is to select the decision rule associated with the lower risk.

There is one additional consideration: the driver may have information other than the color of the light. He may have general information such as the knowledge that high school was just dismissed and it is likely that there are teenagers (who tend to rush the lights) approaching the intersection. He may have more specific information. If there is a policeman at the intersection waving him on, the information from the traffic signal can be ignored

CONFIDENTIAL

where p(ϑ) is the probability that the contact ...

(that is, No. 4 can be used). If a private citizen has undertaken to help by signaling the driver to proceed, the additional information is also ambiguous; however, the risk for No. 3 or No. 4 is certainly reduced.

This may appear to be a complicated way of describing a simple situation. It has two advantages. It is explicit and comprehensive, guaranteeing that every relevant factor can be introduced and weighed. In addition, it is structured in a way which makes it possible to use various analytical tools such as probability theory and utility theory.

The analogies between the example and the ASW classification situation are probably clear. In the latter case, the OBSERVATIONS are signals seen or heard on the various displays and the ACTIONS are "attack" and "ignore" or "report" and "ignore, " etc. (depending upon specific circumstances). The STATES are "contact is a sub" and "contact is a non-sub." The undesirable events are failure to engage a sub and the waste of energy, time, or ammunition on a non-sub.

2.5 The Value of the Information System

In setting up the analysis in mathematical form, it is postulated that all of the relevant concepts can be represented as numerical magnitudes. Many of these magnitudes are probabilities. The ROC provides two very important sets of probabilities: for each possible decision rule it provides the probability that "engage" will be selected when the contact is a non-sub and the probability that "ignore" will be selected when the contact is a sub (assuming for the moment that these are the appropriate actions). In the simple example above, an ROC for the traffic signal would provide the probability that GO is selected when there is cross traffic for Decision Rules 1, 2, 3, and 4, and the probability that STOP is selected when there is no cross traffic for Decision Rules 1, 2, 3, and 4. A second important class of numbers is the "utility losses" associated with the undesirable events. In the ASW analysis, we measure these losses in terms of escorted ships sunk, and estimate these magnitudes for various significant situations.

From the ROC and the utility losses, we calculate the risk (or average loss) for each decision rule and each possible value of the probability that the contact is a submarine. For each value of the probability that the contact is a submarine we determine the lowest risk and the associated decision rule. This is the key step in the analysis. From the utility losses alone we then calculate the risk for each possible value of the probability that the contact is a submarine; that is, the risk when there is no system. (This is equivalent to calculating expected losses at the intersections when the signal is

CONFIDENTIAL

not lit.) This second risk is always greater than or equal to the minimum risk using the system. The difference between these two risks is called the specific value of the system. This specific value is different for each possible probability that the contact is a submarine. The value of the system is measured in terms of the reduction in risk achieved using the system.

The role of the ideal decision maker is to select the optimal decision rule; that is, the decision rule associated with the smallest risk. In general, the identity of the optimal decision rule changes as the probability that the contact is a submarine changes.

In addition to the procedure just described, we compute measures which are independent of the probability that the contact is a submarine. If all of the values of minimum risk are examined, it will always be found that there is one which is largest. This maximum of the minimums is called the "maximin risk." Also there is a decision rule, called the "maximin decision rule" which guarantees that when it is used the risk will not be greater than the maximin value regardless of the probability that the contact is a submarine. An ideal decision maker would select this maximin rule if he knew nothing about the probability that the contact is a sub other than the information derived from the system itself. There is an analogous maximin rule for the case in which there is no system, and the associated maximin risk is greater than or equal to the maximin risk when the system is used. The difference between these two maximin risks is called the general value of the system. It is independent of the probability that the contact is a submarine.

2.6 Applications of the Analysis

(C) Useful conclusions are derived from the analysis by applying it to specific situations. The particular situation we selected is one in which the ASW system is on a vessel assigned to a convoy screen. The losses are measured in terms of ships sunk. Conservative values are selected; that is, it can be expected that in actual situations the losses would certainly be higher than those assumed. In addition, conservative values are selected for sweep rates and contact rates. Using these, the system value is converted to losses per mile of convoy duty. On the basis of the replacement value of the ships lost, the value is further converted into dollars saved per mile. For an actual example of a fairly high performance system, this number turned out to be $480 per mile of convoy duty.

(C) We then analyzed the ROCs obtained at DRL for naive and expert operators (as has been mentioned). Again, on an ultraconservative basis, it was found that the degree of improvement could be evaluated

CONFIDENTIAL

at $96 per mile of convoy duty. This would more than pay for any conceivable investment in personnel training, improved equipment, etc., designed to raise men from the naive to the expert level. This is the key result of the sensitivity analysis for it shows that improved training of classification operators is certainly worth doing: it more than pays for itself in improved overall system performance.

(C) We also examined the relationship between the theoretical and actual procedures. It was determined that there are practical ways of obtaining results close to the theoretical optimals. This requires changes in present procedures and a great deal of future work in developing the details of the modified procedures. The most striking single alteration is to change the sonar operator's output from a choice between "possible sub" and "non-sub" to a range of outputs indicating his subjective estimate of the probability that the contact is either a sub or a non-sub. This recommendation can be extended to the CO when he merely reports a classification decision to the OTC or screen commander and so on up the line to the decision maker who selects the terminal action.

CONFIDENTIAL

3. ANALYSIS

3.1 ASW Classification Systems

(C) An ASW classification system is the combination of equipment, people, and procedural rationale used in carrying out the sequence of operations which begins after detection and which ends with the selection of either an action which is appropriate when the sonar contact is a submarine or an action which is appropriate when the contact is not a submarine. The nature of the terminal actions is an important difference between systems satisfying this definition. In some instances (e.g., sonarmen report to the CO, the CO's report to OTC), the actions may be "call the contact a 'possible sub'" and "call the target a 'non-sub'." In other instances, the available actions may be "attack" and "do not attack." At any instant, there are only two alternative terminal actions since the specific nature of the appropriate terminal action is decided beforehand; for example, if "attack" is an alternative, the specific tactics and weapons to be used are already known. At any instant there is usually a third alternative available; namely, to collect more information before selecting a terminal action. Selection of this alternative increases the operations for which the classification system is used, and providing for this alternative usually complicates the system.

(C) Figure 1 illustrates these concepts. The sequence of operations starts with the reception of acoustic signals from the contact. These are converted into electrical signals by the receiver (which includes the transducer). Usually some information is lost (or deliberately discarded) in this process. In turn, the electrical signals are converted into audio and visual signals by the displays. Again, the information content is usually reduced. At first, the receiver and displays are used in detection and in this mode they are not part of the classification system as described above. Subsequently, they are essential elements of the system. Further, there may be receiver-display combinations which are used only for classification. In the system shown, the next element is "man." This is frequently a team of men. Man carries out three essential functions. He looks at the displays to check for the presence or absence of indications that the contact is (or is not) a sub. This process can be described formally as the conversion of the display outputs to clues. Man then selects an action based on these clues, using a classification rationale which we shall characterize later as a "decision rule." Finally, man can control and adjust the equipment to enhance or eliminate various kinds of information; that is, to change the information losses noted above.

CONFIDENTIAL

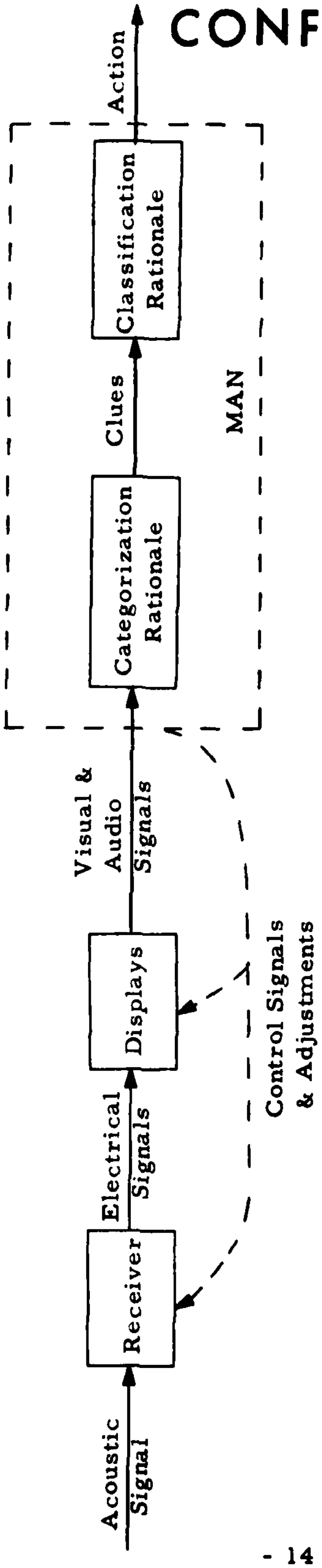

Figure 1. An ASW classification system. (C)

CONFIDENTIAL

(C) There are many significant variations of this system. Man can, and usually does, combine the categorization and classification rationales into some inseparable joint rationale. Alternately, some or all of the classification rationale may be carried out with automata, such as HHIP or MITEC.[2] Another variation is to introduce devices which go directly from the electrical signals to clues, or even to actions. These devices, sometimes in direct imitation of man, may learn or otherwise adapt to the varying circumstances in which the system is used. These variations all illustrate the point that the "system," as defined above, is the collection of entities used in implementing the operations between detection and action selection.

3.2 Receiver Operating Characteristics

The result of using the classification system is one of four possible joint events: (a, s), $(\tilde{a}, s)$, $(a, \tilde{s})$, $(\tilde{a}, \tilde{s})$; where

- s designates that the contact is a submarine.
- $\tilde{s}$ (frequently verbalized "not s") designates that the contact is a non-submarine.
- a designates that the selected action is appropriate when the contact is a submarine.
- $\tilde{a}$ designates that the action selected is appropriate when the contact is a non-submarine.

If the contact is a submarine, the first two events are mutually exclusive since the terminal action must be either a or $\tilde{a}$.

Let

$$y = p(a \mid s) = 1 - p(\tilde{a} \mid s) \tag{1}$$

and

$$x = p(a \mid \tilde{s}) = 1 - p(\tilde{a} \mid \tilde{s}) \tag{2}$$

where $p(\alpha \mid \sigma)$, $\alpha = \{a, \tilde{a}\}$, $\sigma = \{s, \tilde{s}\}$, is the conditional probability that the selected action is α when the contact is σ. The probability, y, is normally called the "hit probability"* and x is normally called the "false action probability."

*This standard terminology, derived from radar technology, refers only to the decision process and not to any subsequent weapon operations.

CONFIDENTIAL

For any given system, there is an additional functional relationship:

$$y = y(x) \tag{3}$$

normally called the "receiver operating characteristic" or "ROC."[2] The ROC describes the inherent capability of the system to distinguish between contacts which are subs and contacts which are non-subs. The significance of the ROC is explained using the representation of a classification system shown in Figure 2.

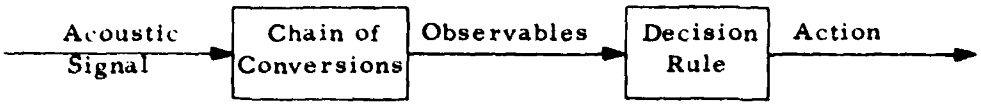

Figure 2. Representation of a classification system.

Every system converts the incoming acoustic signal into a penultimate form called an "observable." (In the system illustrated in Figure 1, the observables are labeled "clues.") The final step is the use of the observables to select an action. This step can be characterized as the use of a "decision rule" which associates one of the alternative actions with each of the possible observables. Let $c = c_1, c_2, \ldots, c_m$ represent the possible observables (or clue combinations). Then the role of the decision rule is described formally by

$$d_i(c) = a \tag{4}$$

that is, the decision rule can be regarded as a function which generates an action once the observable, c, is specified. In general, there are 2^m such rules when there are m possible observables. Figure 3 shows the eight possible rules when $m = 3$; for example,

$$d_3(c) = \begin{cases} a; \text{ when } c = c_1 \text{ or } c_3 \\ \tilde{a}; \text{ when } c = c_2 \end{cases}$$

CONFIDENTIAL

Rule	a	$\tilde{a}$
1	c_1, c_2, c_3	---
2	c_1, c_2	c_3
3	c_1, c_3	c_2
4	c_2, c_3	c_1
5	c_1	c_2, c_3
6	c_2	c_1, c_3
7	c_3	c_1, c_2
8	---	c_1, c_2, c_3

Figure 3. Possible decision rules for three observables.

It can be shown that for the cases of interest here, half of these decision rules can be ignored. When there are only two actions, the observables can be arranged in an order, say c_1, c_2, c_3, such that all of the significant decision rules are generated by shifting through the sequence as shown in Figure 4.

Rule	a	$\tilde{a}$
1'	c_1, c_2, c_3	---
2'	c_1, c_2	c_3
3'	c_1	c_2, c_3
4'	---	c_1, c_2, c_3

Figure 4. Significant decision rules for three observables.

For any particular system there is a hit probability and a false action probability associated with each decision rule. Each point of the ROC represents the pair of these probabilities associated with one of the significant decision rules. When there are m observables, the ROC function, $y(x)$, is not continuous but discrete; and it can be expressed as the set of pairs of numbers, $(x_1, y_1), (x_2, y_2), \ldots, (x_{m+1}, y_{m+1})$. When the ROC function is

CONFIDENTIAL

plotted, it is usually shown as a continuous curve (the "ROC curve" normally pronounced "rock curve"). The general appearance of these curves is shown in Figure 5.[4] A typical association of points on the curves with particular decision rules is also shown. The straight line from (0, 0) to (1, 1) represents a system in which the ability to discriminate between s and $\tilde{s}$ *is no* better than that which can be obtained by pure chance (e.g., by tossing a coin). Any curve which arcs up from this line represents a system which does better than chance.

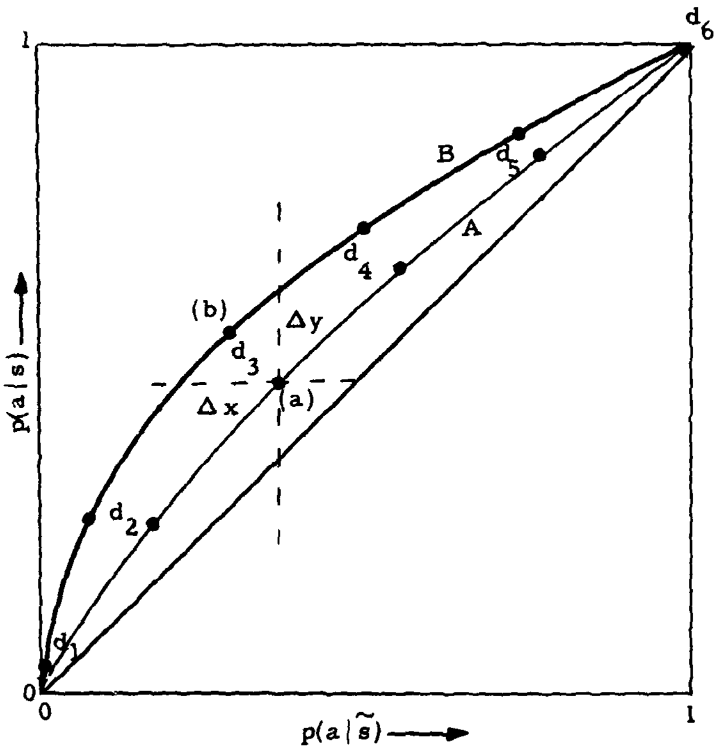

Figure 5. Typical ROC curves.

CONFIDENTIAL

3.3 The ROC as a Measure of Effectiveness

(C) The ROC has been widely accepted as a measure of ASW classification system effectiveness which is very useful in the research and development of new and improved equipment. As a general measure of system effectiveness, this function has strengths and weaknesses. The strengths are illustrated by comparing the systems represented by Curves A and B. For any particular value of the false action probability, System B operates at a higher hit probability than System A: this difference is indicated by Δy in the figure. Analogously, for any given hit probability, System B operates at a lower false action probability than System A: the difference is indicated by Δx in the figure.* It follows that a system designer can examine the ROC curves to determine which of two systems is better. Further, he gets a general indication of the magnitude of the difference--"general" because the difference between the curves decreases with proximity to (0, 0) and (1, 1). The advantage to the designer is that he can compare systems without specific consideration of the decision rules which must be selected to fit the circumstances (as shall be shown subsequently): he can advocate the use of System B rather than System A on the basis that, in all possible circumstances, it is usually better and never worse than System A.

The weakness of the ROC as a general measure of system effectiveness may already be apparent in the preceding statements. Sooner or later it becomes necessary to face the question of how much better is System B and to undertake an analysis of the selection of decision rules. In system design this question usually arises when cost-effectiveness is considered, since it is then necessary to determine whether the improvement is worth the cost. In this study, the question arises because we are primarily interested in operator performance, and all of the mistakes the operator may make have as a consequence the selection of the wrong decision rule. Consequently, this question is analyzed in the following sections. Subsequently

---

*This argument may not seem completely convincing if it is noted that the physical significance of a point on the ROC curve, lying between two points representing different decision rules, has not been explained; for example, the point between d_3 and d_4 in the illustration. The intermediate points are associated with "mixed decision rules"; however, these are not normally used in practice because points like d_3 and d_4 are usually close enough to be good approximations to the mixed rule. In the illustrated instance, when d_3 is used, the advantage of System B over System A is the difference between operating at point (b) rather than (a), and in this case a higher hit probability and a lower false action probability are achieved simultaneously.

CONFIDENTIAL

in discussing operator performance, we adopt the attitude that any particular ASW system is described by ROC, and we use actual ROCs as illustrative inputs to our analyses. At this stage in our study, we are not particularly concerned with the problem of actually determining the ROCs. Later, in order to measure operator performance, we shall have to investigate this process in detail.

3.4 Expected Losses

(C) The simplest class of cases is that in which only the consequences of erroneous action are significant. These consequences are called "utility losses" or more simply just "losses." Frequently these losses can be measured in tangible terms such as ships sunk, weapons wasted, or lost time. In game situations, such as exercises, the utilities may be intangible: for instance, they may be measured in points scored according to some rules. Such utilities are important to the individuals involved because of a conscientious desire to perform well, pride, and possible effect on career advancement. They are important to the Navy because of the indication of the capabilities to be expected in combat. Utilities may also be probabilistic, in that they may involve several alternate events, each with a different utility and a different probability of occurrence. There is an extensive body of theoretical and practical knowledge which indicates that it is always possible to define sensible, useful utilities in the face of these and many other contingencies. We accept the results of this "utility theory" without further discussion. (For details see Refs. 5, 6, 7, 8.)

(C) As a prototype of situations in which only losses are significant, consider the case in which the ASW classification system is on a ship assigned to a screen protecting a group of ships--a convoy, or perhaps a landing support force operating off the beachhead. Under these circumstances, the ASW operation is successful as long as none of the protected ships are lost. (The ASW ship itself may well be considered expendable.) Clearly, this is a case in which only losses, measured in terms of ships sunk, are significant.

Losses may occur in spite of the best efforts of the protective force, but these are not significant in the analysis. The significant operational variation is the selection of the appropriate action. Losses which may occur when the appropriate action is selected do not, in any way, change the identification of the action as "appropriate"; if a better action is available, it is the one which will be so labeled. In this case, therefore, the analysis is reduced to a consideration of the losses which result from the selection of the inappropriate action; that is, the losses associated with

CONFIDENTIAL

the terminal events $(a, \tilde{s})$ and $(\tilde{a}, s)$. This is illustrated by the tableau in Figure 6-A which indicates a loss, L_1, when the contact is a sub and the inappropriate action (usually no action at all) is selected, and a loss, L_2, when the contact is a non-sub and the inappropriate action (report to higher echelon, attack, etc.) is selected.

The normalized form of the tableau, shown in Figure 6-B, is obtained by dividing each element of the explicit form by L_1, and setting $\lambda = L_2/L_1$. It is proved in general that this does not alter any of the results of interest, and there is a considerable reduction in algebraic complexity.[9]

A. Explicit Form

		Nature of the Target	
		Sub (s)	Non-Sub ($\tilde{s}$)
Action Selected	a	0	L_2
	$\tilde{a}$	L_1	0

B. Normalized Form*

		Nature of the Target	
		Sub (s)	Non-Sub ($\tilde{s}$)
Action Selected	a	0	λ
	$\tilde{a}$	1	0

Figure 6. Losses associated with inappropriate actions.

(C) The magnitude of L_1 (the loss when a sub is misclassified as a non-sub) depends upon several contingencies as follows:

* When a specific name is required, the magnitude of the entries in this tableau are usually referred to as "utiles."

CONFIDENTIAL

CONFIDENTIAL

1. The appropriate action may be initiated by another ASW ship in the screen.

2. The submarine may not attack.

3. The submarine may attack ineffectively by selecting relatively unimportant targets, as a result of malfunctions in his weapon system, or even ineptitude.

Each of these contingencies can be broken down further. As a result, a very large number of parameters must be specified before it is possible to specify the events and probabilities which must be known to calculate L_1. However, this complex set of possibilities is highly redundant, since all cases must reduce eventually to a statement of an expected number of ships lost (L_1), and the number of possible values of L_1 is very limited (as a first guess to a number less than or equal to five).

(C) The calculation of L_2 is similarly complex. The consequences of selecting action "a" when the contact is a non-sub is a diminution in defensive capability. This may be a temporary condition; for example, an SAU may be dispatched on a "wild goose chase" and during this interval, it is not effective against some sub which may pose an immediate threat. The diminution may be longer lasting as occurs when there is a limited supply of weapons and some are wasted on a non-sub target. Again, despite the complexity of possible situations, the number of end results is limited since L_2 is almost certainly less than L_1. With ingenuity, it is possible to delineate situations in which risking an attack by a sub leads to a smaller loss than attacking a non-sub, but these are so unique that they will almost certainly never arise in practice. Thus, $L_2 \leq L_1$ (and $\lambda \leq 1$) and L_2 is at least as limited in range as L_1. Accordingly, we continue the analysis in the expectation that while the task may be tedious, reasonable estimates of L_1 and L_2 can be obtained when needed. Later we shall also "discover" that the analysis produces useful results even when the loss estimates are not very accurate.

3.5 Risk

Given a system, and hence a ROC, there is a specific false action probability associated with each significant decision rule (or point on the ROC curve). This probability is designated by $p_i(a|\tilde{s})$. The probability of selecting the "false action," when decision rule d_i is used, is

$$p_i(a, \tilde{s}) = p_i(a|\tilde{s})\, p(\tilde{s}) \qquad (5)$$

where $p(\widetilde{s})$ is the probability that the contact is a non-sub. At the moment, we merely define $p(\widetilde{s})$ arbitrarily; later we will interpret this as either the "real" probability (e.g., known to some superior intelligence) or as some a priori estimate of the probability. Using the definition of x implicit in Eq. (2) and letting $p = p(s) = 1 - p(\widetilde{s})$, Eq. (5) becomes:

$$p_i(a, \widetilde{s}) = x_i(1 - p) \qquad (6)$$

and this is the probability, using the normalized tableau, that a loss of λ will result from the use of the given system and the decision rule d_i. From Eq. (1) the probability that $\widetilde{a}$ will be selected when the contact is a sub, using the same system and d_i, is

$$p_i(\widetilde{a}, s) = (1-y_i)p \qquad (7)$$

and in this case a loss of unity is sustained. It follows that the expected value of the loss when d_i is used is

$$R(i, p, \lambda) = x_i(1-p)\lambda + (1-y_i)p \qquad (8)$$

The symbol R is selected because this expected loss is usually referred to as "risk." As indicated, the risk with any given system will change with the decision rule, d_i; the probability that the contact is a sub, p; and the loss parameter, λ. The risk will also change when the system (i.e., the ROC) is changed. Since λ is a constant, and x_i and y_i are constant for a given decision rule, a plot of $R(i, p, \lambda)$ against p for a given i and λ is a straight line. Figure 7 shows a hypothetical example in which there are three decision rules (i.e., i = 1, 2, 3).

If there are m possible values of i, then there are m values of $R(i, p, \lambda)$ for a given λ and any particular value of p. For any particular value of p, at least one of these risks is the smallest. Let i^* designate a value of i associated with this smallest risk. In Figure 7, $i^* = 1$ for all values of p between 0 and p_b; $i^* = 2$ for p between p_b and p_c; and $i^* = 3$ for p between p_c and unity. The set of all of these minimum risks (i.e., for each value of p) is designated by $R^*(p, \lambda)$ or $R(i^*, p, \lambda)$; and the operation which generates $R^*(p, \lambda)$ is indicated explicitly by

$$R^*(p, \lambda) = \min_i R(i, p, \lambda) \qquad (9)$$

CONFIDENTIAL

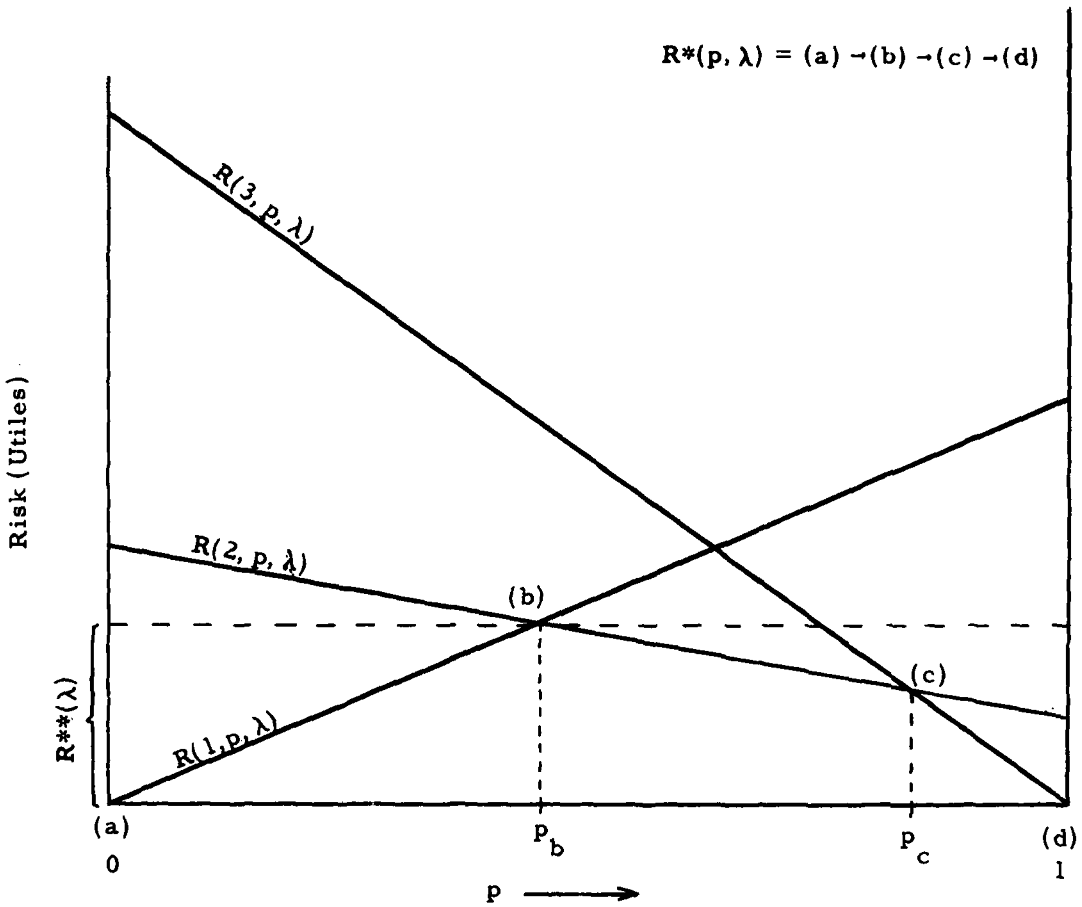

Figure 7. Risk, minimum risk, and maximum risk.

which merely means that $R^*(p, \lambda)$ is formed by examining $R(i, p, \lambda)$ for a given value of p and every value of i, selecting the smallest of these values, and then repeating the process for the next value of p.

A physical interpretation of this result is that a decision maker who knows the ROC, the value of the loss parameter, λ, and the probability, p, can select the decision rule associated with the smallest risk. When he does this, the risk is $R^*(p, \lambda)$. We call this hypothetical individual the "ideal decision maker." A real decision maker will usually know very little (if anything) about p and λ. Further, he may arrive at his decision via a procedure which in no way resembles this analytical approach. However, regardless of the rationale the real decision maker employs, or the extent of his knowledge, the best the real decision maker can do is to select the decision rule, which we have designated by i*, with the associated risk (which he may not know) $R^*(p, \lambda)$. If he selects any other decision rule, the risk will usually be greater and will never be smaller than $R^*(p, \lambda)$. Thus, the ideal decision maker represents the limit that the practical decision maker can achieve.

3.6 <u>Maximin Risk</u>

The plot of $R^*(p, \lambda)$ in Figure 7 (that is, the plot through the points (a), (b), (c), and (d)) is typical of all possible instances in that it is continuous and has a maximum value for some value of p between zero and unity. In Figure 7, this maximum occurs at the point (b). This largest value of $R^*(p, \lambda)$ is designated by $R^{**}(\lambda)$. The operation is indicated formally by

$$R^{**}(\lambda) = \max_p R^*(p, \lambda) = \max_p \min_i R(i, p, \lambda) \qquad (10)$$

Because of the form of the last expression, $R^{**}(\lambda)$ is called the "maximin risk."

(C) There is a decision rule which results in a risk of $R^{**}(\lambda)$ for all possible values of p: it is called the "maximin decision rule" and designated by the index i **. For the case illustrated in Figure 7, this rule would specify using the decision rule, d_1, with a probability Φ, and the rule d_2 with a probability $1-\Phi$, where Φ is chosen as the value which results in an expected value of the combination of the risks $R(1, p, \lambda)$ and $R(2, p, \lambda)$ lying along the dotted horizontal line passing through (b). Using such a rule in practice would be quite cumbersome. Fortunately, there is usually a simple decision rule which produces a result very close to the theoretical ideal. In Figure 7, the closest approximation is to use the rule d_2 (that is, i**=2). In this case, the largest risk occurs when p = 0; that is, at the point at which $R(2, p, \lambda)$ intersects the vertical axis through p = 0. This value is not very much

CONFIDENTIAL

larger than $R^{**}(\lambda)$. In actual cases the approximation will invariably be closer. Figure 8 shows the risks for an actual ROC with ten significant decision rules when $\lambda = 1$. It will be noted that $R(6, p, \lambda)$ passes through the maximin point and is almost horizontal. As a consequence, d_6 is a very close approximation to the maximin decision rule. Further, d_5 is just as good as d_6. The increase in maximum risk using d_5 or d_6 instead of the true maximin mixture is .03 utiles. In Figure 9, the curve for $R^*(p, \lambda)$, when $\lambda = 1$, is repeated and curves for other values of λ are shown. As might be expected, the risks decrease as λ decreases. A perhaps unexpected characteristic of these curves is that the optimal decision rule varies very little as λ ranges from 0.2 to 1.0: the optimal rule shifts from a mix of 5 and 6 to a mix of 7 and 9 (note: Rule 8 is not used in this instance).

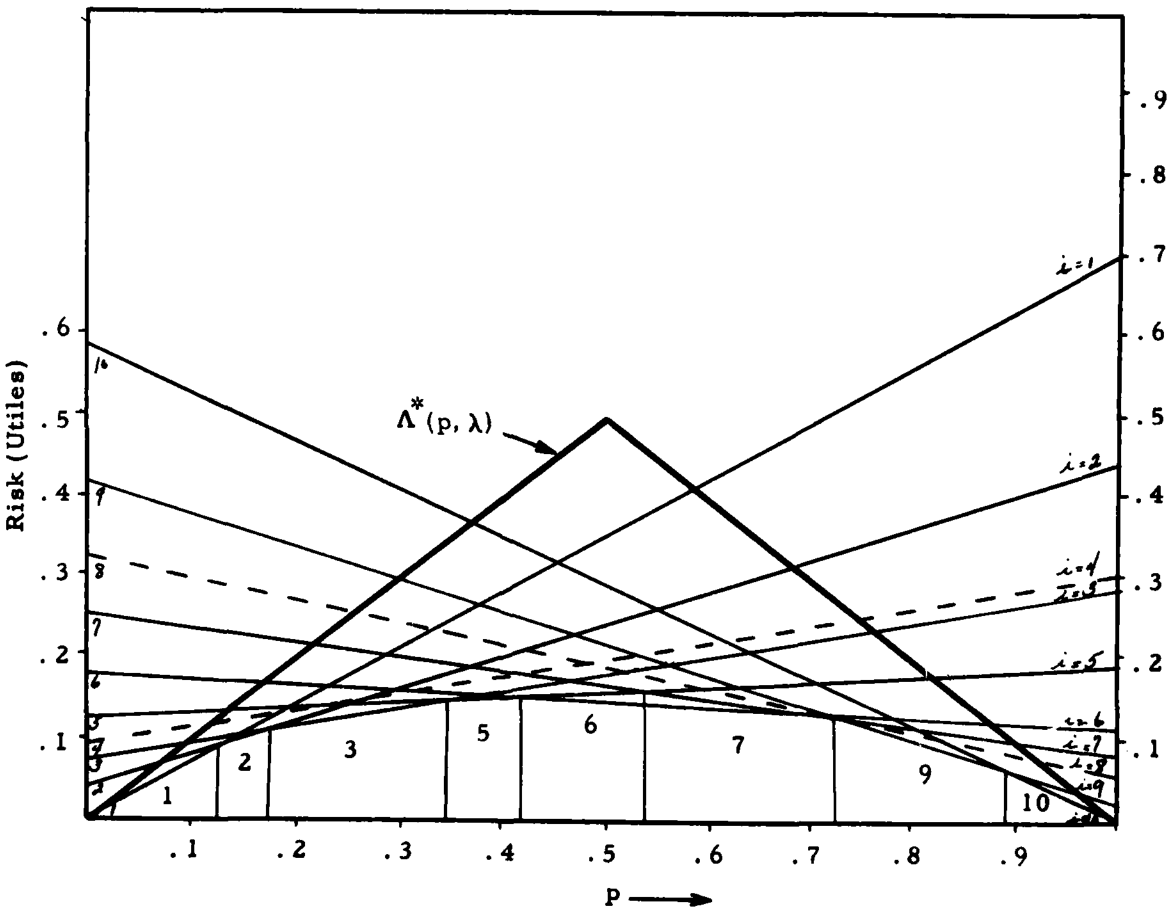

Figure 8. Risks for case with ten decision rules. (based on ROC in Reference 2). (C)

CONFIDENTIAL

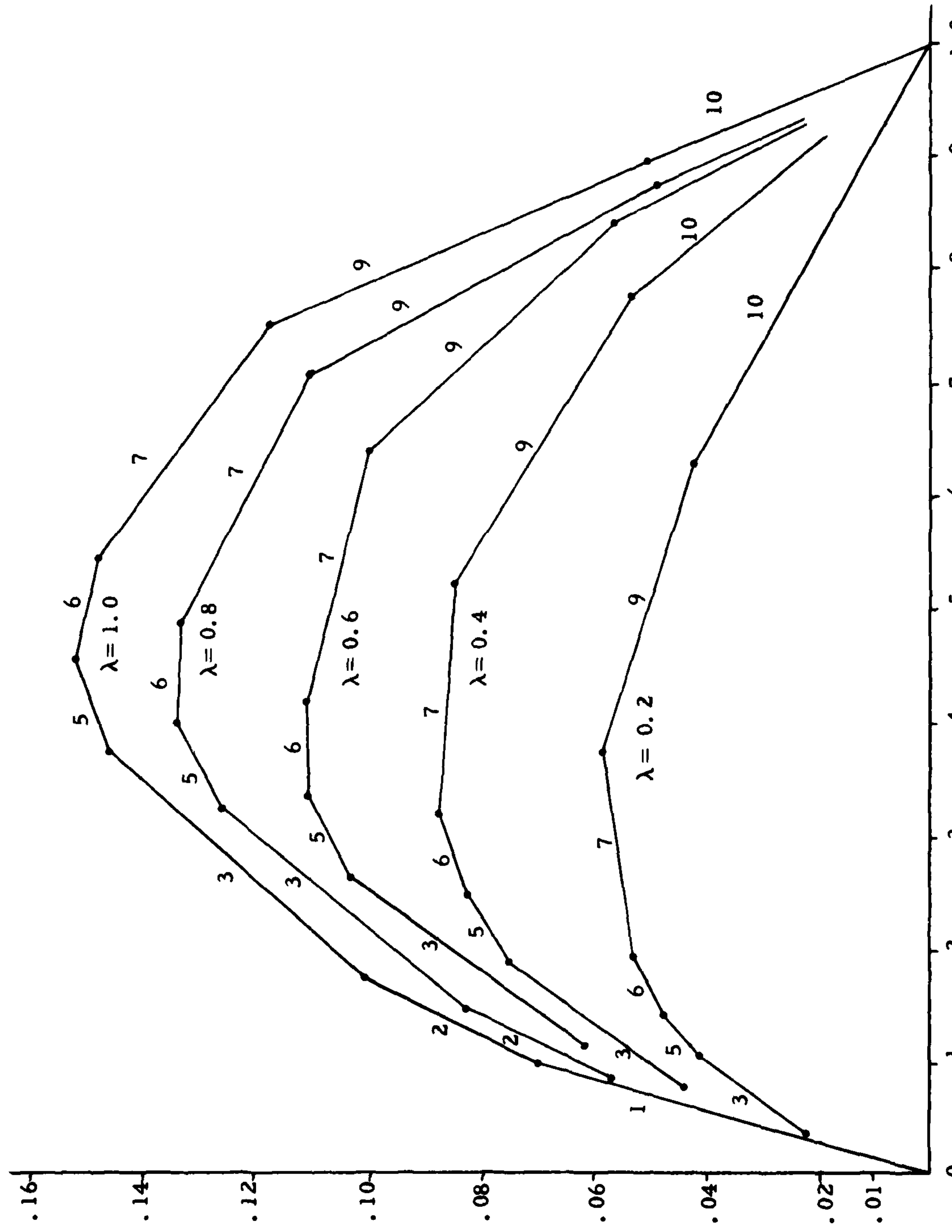

Figure 9. Minimum risks for various losses. (C)

CONFIDENTIAL

3.7 Value of a System

We have already introduced the concept of an ideal decision maker characterized by his knowledge of:

1. the ROC
2. the loss parameter, λ
3. the probability, p, that the contact is a sub (or non-sub)
4. the procedures for determining minimum risk

and we have shown that he can *always select* a decision rule, d_{i*}, which results in the lowest risk, $R^*(p, \lambda)$. Further, we have shown that if the ideal decision maker does not know the third item, the probability p, he can use the maximin decision rule d_{i**} or a close approximation to it and limit his risk to $R^{**}(\lambda)$ regardless of the actual value of p. Finally, we have noted that an actual decision maker, regardless of what he knows or what procedure he uses in arriving at a decision cannot (on the average) do better than the ideal decision maker, but he may do much worse. The ideal decision maker represents the upper limit to what the actual decision maker can accomplish.

Suppose that the ideal decision maker does not have a system and tries to operate without the ROC, or equivalently without the values of x_i, y_i. On the basis of his knowledge of p and λ, he can select the action which will result in the smaller risk. In particular, if he selects action a, the expected loss, or risk, is that which is incurred given that the contact is a non-sub, multiplied by the probability (1−p) that the contact is a non-sub. Formally,

$$\Lambda(a, p, \lambda) = (1-p)\lambda \tag{11}$$

Analogously, if he selects the action $\tilde{a}$, the risk is

$$\Lambda(\tilde{a}, p, \lambda) = p \tag{12}$$

The ideal decision maker can select the action, say a^*, which is associated with the smaller of these two risks. Formally, he can limit his risk to

$$\Lambda^*(p, \lambda) = \Lambda(a^*, p, \lambda) = \min_{a} (a, p, \lambda) \tag{13}$$

As long as the ROC represents a system which does better than chance, this risk is larger than $R^*(p, \lambda)$, determined from Eq. (9). The difference between the risks without and with the system is a measure of the "value" of the system. Formally, we define

CONFIDENTIAL

$$V^*(p, \lambda) = \Lambda^*(p, \lambda) - R^*(p, \lambda) \quad (14)$$

and we call this the "specific value of the system" (because it applies to a specific value of p).

Figure 10 illustrates the determination of $V^*(p, \lambda)$. The $R^*(p, \lambda)$ curve is repeated from Figure 8. $\Lambda^*(p, \lambda)$ is the lower boundary of the two straight lines representing $\Lambda(\alpha, p, \lambda)$ given in Eqs. (11) and (12). The difference between these two is the dashed line marked $V^*(p, \lambda)$.

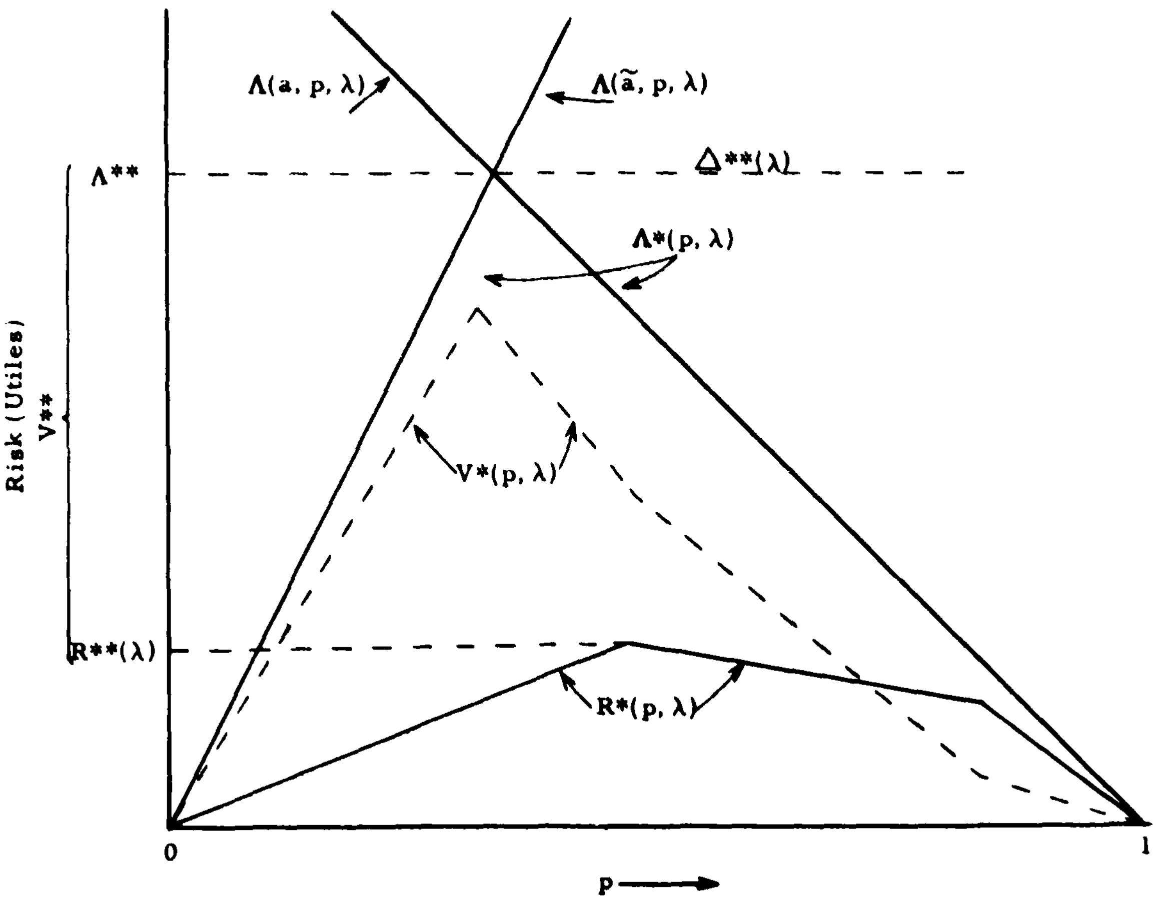

Figure 10. Specific and general values of a system.

CONFIDENTIAL

Some insight into the interpretation of V* may be obtained by deriving an obvious conclusion from the fact that it is zero when $p = 0$ or when $p = 1$. V* is the value of the system to the ideal decision maker. When p is zero or unity, this decision maker is already completely informed about the nature of the contact; that is, if $p = 1$, he knows for sure that the contact is a sub. In this case, the system cannot add to his information and accordingly its value, V*, should be zero. In general, V* increases as the degree of certainty about the classification of the contact decreases.

Some additional characteristics of V* are shown in Figure 11 in which this function is plotted for several different values of the loss parameter, λ, for the same system (ROC) used in Figures 8 and 9. These curves illustrate that the value of the system increases as the consequences of the false action, $(a, \tilde{s})$, became relatively more serious as reflected by the increasing values of λ.

In analogy with the previous analysis, the ideal decision maker can select an action which leads to a maximin expected loss when he operates without the system. This is a "mixed action" in which he uses a with probability Φ and $\tilde{a}$ with probability $(1-\Phi)$. From Eqs. (11) and (12), his expected loss is

$$\Lambda(\Phi, p, \lambda) = (1-p)\lambda\Phi + p(1-\Phi) \qquad (15)$$

and it is readily checked that when

$$\Phi = 1/(1+\lambda) = \Phi^* \qquad (16)$$

Eq. (15) yields

$$\Lambda(\Phi, p, \lambda) = \Lambda^{**}(\lambda) = \lambda/(1+\lambda) \qquad (17)$$

and furthermore that

$$\Lambda^{**}(\lambda) = \max_{p} \min_{\alpha} \Lambda(\alpha, p, \lambda) \qquad (18)$$

so that Λ^{**} is the maximin loss with the property that use of the mixed action Φ^* ensures that the loss will not be greater than this regardless of the actual value of p. When the ideal decision maker does not know p, he can use the system to reduce this expected loss to $R^{**}(\lambda)$. Accordingly, we define

$$V^{**}(\lambda) = \Lambda^{**}(\lambda) - R^{**}(\lambda) \qquad (19)$$

as the "general value of the system."

CONFIDENTIAL

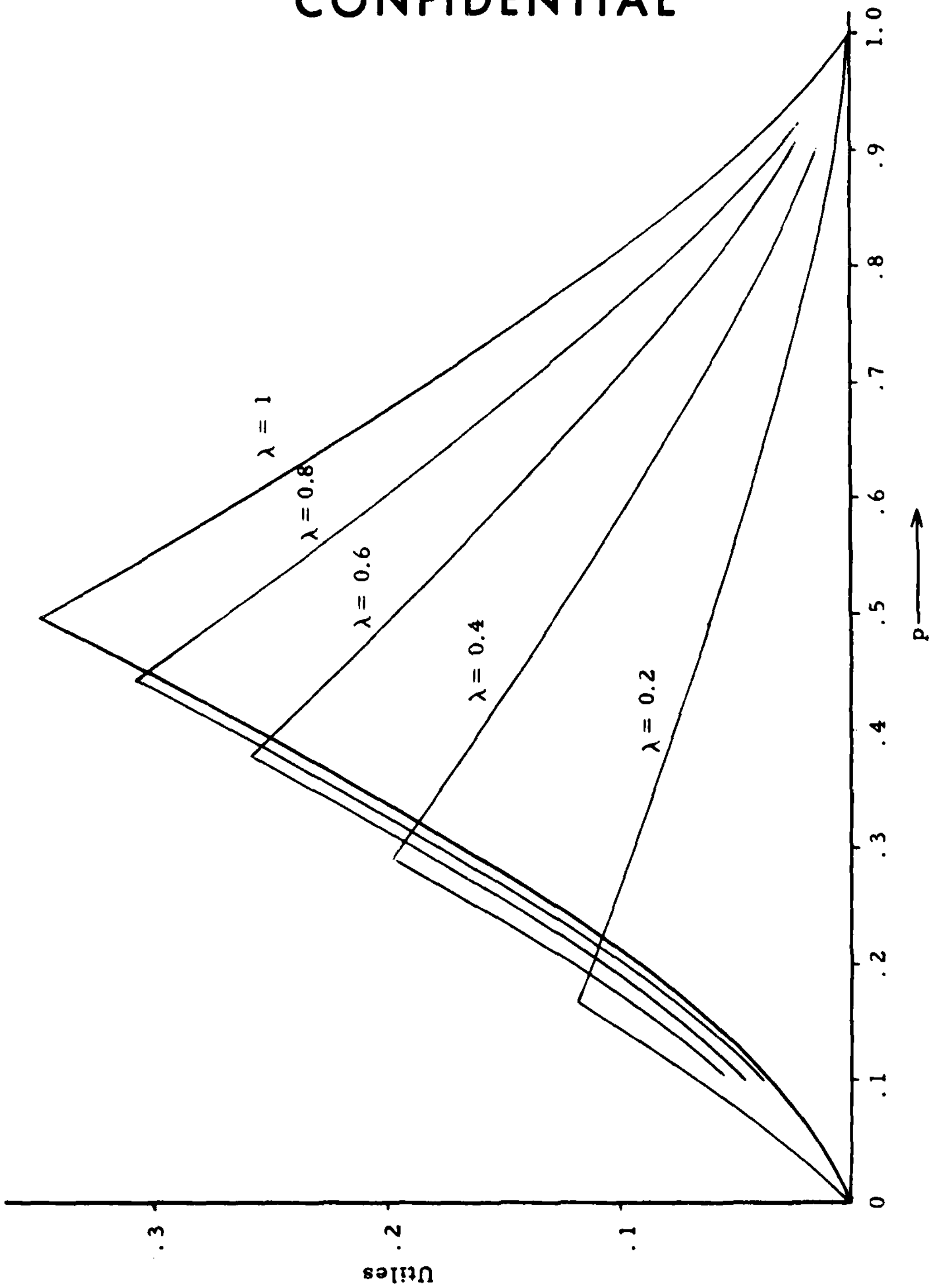

Figure 11. V*(p, λ); specific values for various losses. (C)

Values of V** for the system illustrated in Figures 8, 9, and 11 are shown as System 1 in Figure 12. Two other cases which will be discussed in detail subsequently are also shown.

bility are achieved simultaneously.

- 19 -

CONFIDENTIAL

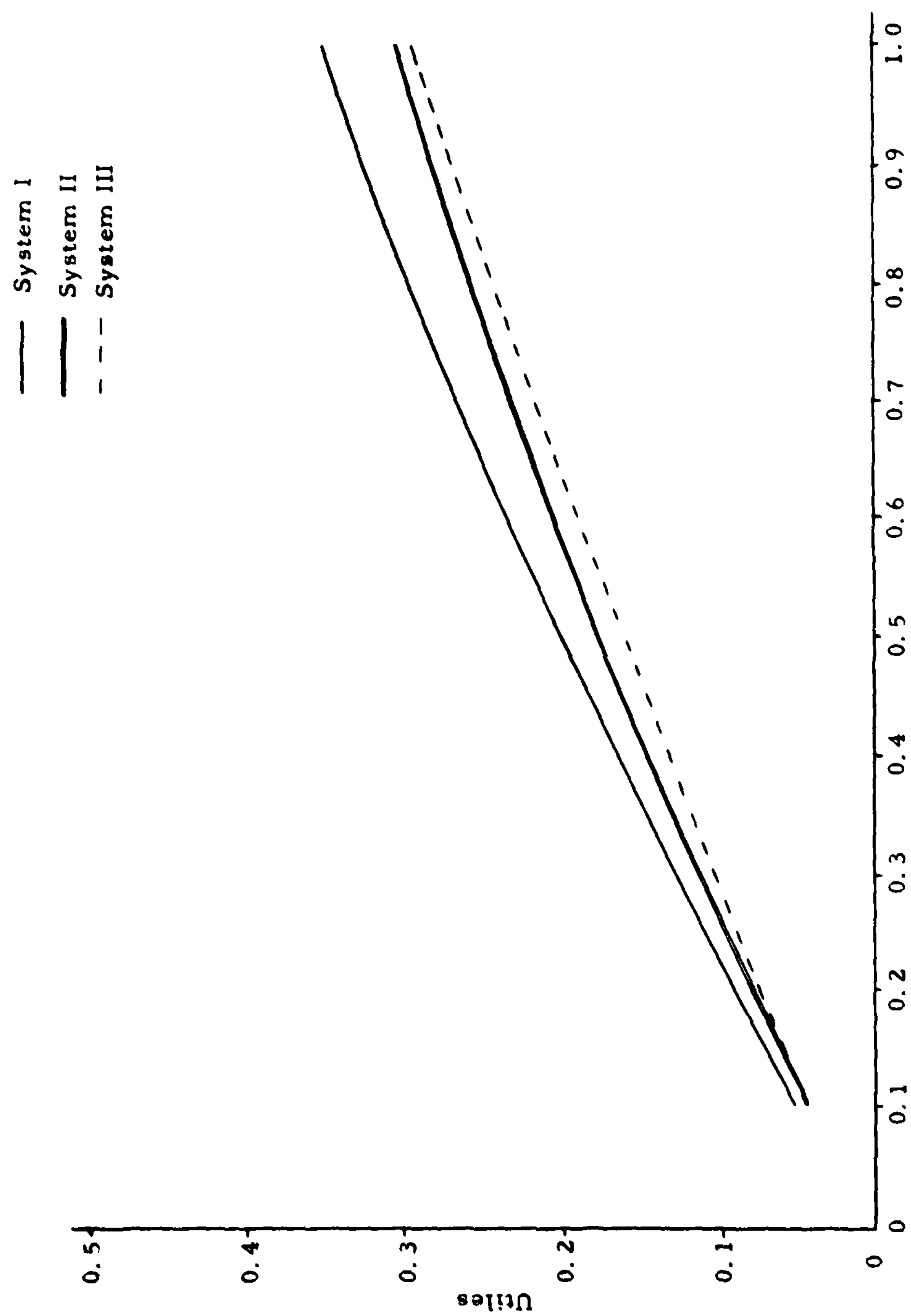

Figure 12. V**(λ): general values of three systems. (C)

4. APPLICATIONS OF THE ANALYSIS TO SYSTEM EVALUATION

4.1 Analysis of Value in Convoy Situations

The analysis developed in the preceding sections can be used to obtain a practical evaluation of ASW classification systems by applying it to investigations of specific circumstances. It is clear that knowledge of the circumstances is important since a system can have no value if it is not used, as may happen if the contingencies it is designed to cope with never arise. In operations as complex as those in which antisubmarine warfare is conducted, there is no simple way of characterizing these circumstances. We are reduced, therefore, to an examination of specific examples which are judged to be reasonably representative and significant.

The first example is a continuation of the one introduced in the preceding section. Suppose that the ASW system is on a vessel engaged in convoy duty. As has been demonstrated in the development of the analysis, in the absence of any prior information about the probability that a contact is a submarine, it is reasonable to use the maximin decision rule. Without the classification system the risk, or expected value of loss, will then be $\Lambda^{**}(\lambda)$ (obtained by selecting the mixed action associated with Φ^*, cf. Eq. (17)). When the system is used, the risk will be $R^{**}(\lambda)$ (or a very close approximation, cf. Eq. (10)). As a consequence, the use of the system in any instance will result in a saving of $V^{**}(\lambda) = \Lambda^{**}(\lambda) - R^{**}(\lambda)$ as indicated by Eq. (19).

A point that frequently requires explication is that the saving of V^{**} is achieved regardless of the actual value of the probability that the contact is a submarine. This follows from the fact that the risk associated with the maximin decision rule is an horizontal line on Figure 7, indicating that the risk is the same for all values of p. Since the expected loss associated with the optimal action, Λ^{**}, when there is no system is an horizontal line on Figure 10, also indicating an independence of p; it follows that V^{**} must be independent of p. The situation is analogous to that encountered in the game of matching pennies. If one player selects his alternative (heads or tails) at random, the probability of a match is 1/2 regardless of what the other player does. For example, the probability of a match is 1/2 even when the other player always selects heads. The analogy can be carried further: if one knows the opponent's pattern (e.g., always plays heads), it is possible to achieve a better result than that obtained from use of the maximin random selection. Similarly, it is possible to capitalize on a knowledge of the probability, p, to obtain a risk lower than V**. In this example it is assumed initially that this cannot be done.

(C) Up to this point the units of measurement of V** have been labeled "utiles" to indicate that its magnitude represents the utility value of an event. In particular, the normalized form of the loss matrix, Figure 6-B, indicates that one utile is the utility loss which occurs when a contact which is actually a submarine is misclassified as a non-submarine. The actual consequence of this event is some expected value* of the number of ships sunk by the enemy. This number, in turn, depends on such factors as whether or not the submarine attacks, the effectiveness of his attack, and the possibility that some other ship in the screen will detect the submarine and take the appropriate action. In this example it is assumed that the circumstances are such that this value is one ship sunk. This is a reasonable expectation in a large variety of circumstances, and a useful starting point since it simplifies some of the interpretation tasks.

(C) Since V** is earned each time the system is used, the total value of the system is NV**, where N is the number of times used. In the convoy situation, we can estimate a lower limit for N and hence for the total value. A "false contact" is a sonar target (that is, a physical entity which reflects the sonar energy) which is not a submarine. A false contact rate can generally be estimated fairly accurately for a given geographical region and season. Let R stand for the number of false contacts per square mile. Further, suppose that the ship searches over a path of width, W, on a cruise of D miles, covering an area of WD square miles. The number of false contacts is WDR. During the cruise, there will be at least this number of contacts: in addition, there will be contacts which are submarines. Thus,

$$N \geqq WDR \tag{20}$$

and

$$NV^{**} \geqq WDRV^{**} \tag{21}$$

The sweep width, W, associated with a single ship is determined primarily by the spacing between escorts. This spacing, in turn, is a function of several variable factors such as effective sonar range, the shape of the screen, and the number of escorts available. Since we are seeking

*"expected value" is used in the technical sense sometimes loosely indicated by the use of "average values."

a lower bound, we examine the case for which W = 1 n.m., a value far below the average. We set R = 1.6 contacts per 1,000 square miles on the basis of expert advice that this is a reasonable approximation to general experience. Under these conditions,

$$NV^{**} \geq (1.6 \times 10^{-3})\, D\, V^{**} \qquad (22)$$

where D is the distance traveled.

(C) As shown in Figure 12, V** decreases as λ decreases. The parameter, λ, is a measure of the loss associated with the false action; that is, misclassifying a non-submarine as a submarine and acting accordingly. One possible action is the expenditure of weapons. The associated loss involves both the direct cost of the wasted weapons and the decreased capability to deal with future threats. A second possible action consists of reporting the contact as a submarine to a higher echelon which then vectors another ship (e.g., a "pouncer") to the target position. In the course of this action, it is possible that the misclassification will be discovered before any weapons are expended. The loss then will consist of the decreased defensive capability during the period when the force was at least partially preoccupied with the non-submarine contact. This is a smaller loss than that of the useless expenditure of weapons. An overly conservative estimate, consistent with the fact that we are establishing a lower bound, is to set λ equal to 0.1. This is interpreted literally as an assumption that the decrease in defensive capability accumulated in the course of ten responses to false classifications of non-subs as subs gives the enemy about as much advantage as one misclassification of a sub as a non-sub. From Figure 12, V** = 0.05, when λ = 0.1. Then follows that

$$NV^{**} \geq (8 \times 10^{-5})\, D \qquad (23)$$

Thus, in 1,000 miles of convoy duty the classification system can be credited with preventing losses at least equivalent to the loss of .08 ships. Since the replacement value of a merchant ship and its cargo is at least $6 million, this result can also be interpreted as earning $480 per mile of convoy duty. The cost of a classification system is only the marginal cost of the equipment and training which must be added to the system required for detection and fire control. It follows that the investment needed to achieve the capability provided by the classification system used in this example is economically justified. *Further, this line of reasoning strongly suggests that even a small increment in classification capability is worth a sizable investment in training and equipment.*

4.2 Convoying, Continued

Recapitulating the arguments in the preceding section, V** represents the utility when the classification system, represented by the ROC used in determining V**, is used by an informed, rational decision maker. On the average, actual decision makers cannot be expected to do as well.

(C) A conservative estimate of V** is derived from two assumptions. The first is that the expected value of the loss when action is selected on the basis of an erroneous classification of a sub as a non-sub is one escorted ship sunk. If this number is higher, as it usually will be, then V** would also be higher. The second assumption is that the loss when an action is selected on the basis of an erroneous classification of a non-sub as a sub is 1/10 of the loss resulting from the sinking one escorted ship. This is a low estimate of the consequences of treating a false alarm as an actual sub.

(C) Since V** is the value each time the system is used, the total value depends upon the number of uses. In the absence of any submarines this will be just the number of false alarms. This rate, estimated conservatively, times V** leads to the conclusion that the classification system studied saves 0.08 convoyed ships for every 1,000 miles of convoy duty. This is enough to pay for the training and equipment devoted strictly to classification many times over.

(C) In addition, we note that the situation itself is a limiting one, since we have assigned the ASW ship (and its classification subsystem) a very passive role in the convoy: we have assumed essentially that it holds its position in the screen and reports its classification decision to higher echelons. Suppose, however, that the appropriate action is to initiate an attack when it is decided that the contact is a sub: in this case, the value of V** is larger. This is shown in the following argument.

The simple loss matrix used in the preceding analysis is not adequate for the analysis of the active assignment situation since no mechanism for representing the possible gain when a submarine is sunk is available. Ideally, the analysis should be revised somewhat to accommodate a "utility function" describing the gains and losses shown in Figure 13. In particular, when a contact which is actually a sub is attacked, the utility is a gain, G. The losses, L_1 and L_2, are those already defined, but the associated utilities are explicitly negative. The utility of not attacking a non-sub is assigned the value zero arbitrarily.

	Sub	Non-Sub
a	G	$-L_2$
$\tilde{a}$	$-L_1$	0

Figure 13. *Utilities for possible terminal events.*

(C) We plan to study this more general formulation in the very near future. At the moment we can derive a preliminary result by introducing the concept of "regret." Regret is the utility loss associated with not selecting the best possible action in each possible terminal situation. Algebraically, regrets are obtained from the tableau in Figure 13 by subtracting the largest number in each column from all other elements in the column. This produces the regret tableau shown in Figure 14, since G is the largest value in the first column and 0 is the largest value in the second. Finally, the regret tableau can be converted into a normalized form by dividing each of its elements by $-(G + L_1)$, producing the tableau shown in Figure 15.

	Sub	Non-Sub
a	0	$-L_2$
$\tilde{a}$	$-(G+L_1)$	0

Figure 14. Regrets for possible terminal events.

	Sub	Non-Sub
a	0	$\frac{L_2}{G+L_1}$
$\tilde{a}$	1	0

Figure 15. Normalized regrets.

Let

$$\lambda' = L_2/(G + L_1) \quad (24)$$

Then the analysis can proceed as before using λ' instead of λ. It is clear that λ' is less than λ since G and L_1 are always positive numbers. On this basis alone, it would be concluded that V** is lower when the possibility of active assignment is introduced. However, there is one additional consideration. What we have done can be explained as introducing an increased value of L_1; namely,

$$L_1' = L_1 + G \quad (25)$$

This is reasonable because the loss associated with *failure* to attack a sub is now measured not only in terms of losses in escort vessels but, in addition, in terms of the lost opportunity to sink an hostile submarine. *However*, we have ignored the fact that L_2 must also be modified. When the SAU takes action appropriate to the classification of the contact as a submarine, it almost invariably attacks. It moves out of what would normally be its optimal position, weapons are expended, and large amounts of noise are introduced into the acoustic environment.

(C) If the contact is not a submarine, the losses are much higher than those consequent upon the mere reporting of an erroneous classification by a ship holding its position in the screen. It follows that it is necessary to use $L_2' \geqq L_2$, and

$$\lambda' = L_2'/L_1' \quad (26)$$

(C) The gain, G, associated with attacking a submarine is just the kill probability, P_K, for the sequence of actions starting after classification (not P_K for the weapon alone after localization is completed). In addition, G must be expressed in the same units as L_1 or the indicated addition will be meaningless. To convert P_K from "subs sunk" to "escorted ships sunk," we multiply by the "exchange ratio," ρ, a number expressing the fact that one submarine is "worth" ρ escorted ships. Thus,

$$G = \rho P_K \quad (27)$$

On the basis of replacement costs only, $\rho = 5$. (Considering strategic and tactical consequences may result in exchange ratios as high as 20.) A conservative estimate of P_K, as defined above, is about 0.05. Previously, we assumed that $L_1 = 1$. Thus,

$$L_1' = G + L_1 = 0.25 + 1.00 = 1.25 \qquad (28)$$

Previously we assumed that $L_2 = 0.1$. If $L_2' = .125$, then $\lambda' = \lambda$. On the basis of the preceding discussion, it is estimated that the increase in L_2 will certainly be greater than this increase from 0.1 to 0.125, so that λ' will certainly be greater than λ and V** for the active assignment will be larger than that calculated for the ship in the screen. It follows that the original analysis represents the limiting situation.

4.3 Effectiveness of Operator Training

(C) Work currently being carried out at DRL has produced ROCs representative of the results that can be achieved using systems with "expert," "average," and "naive" sonar operators.[10] It is not suggested, either by DRL or here, that these functions are necessarily accurate indications of the results which might be obtained at sea. However, the experiments are realistic, and the differences in operator skills correlate with experience and other independent indications of capability. It follows that for the order of magnitude estimates required here, it is reasonable to accept these data as indications of the differences in skills if not of the absolute level of skills.

(C) Analysis of the ROCs to determine the values of V**(λ) produced the results shown in Figure 16. It is immediately apparent that there is a significant difference between the results achieved with the three levels of operator skill. It follows that this is at least a first order indication of the increased value which might be achieved as a result of personnel selection, training and experience. The results are even more instructive when they are replotted as shown in Figure 17. For low values of λ (equivalent to low false action penalties), operator skill can increase the system value from zero (or very close to zero) to a significant value. For example, when λ 0.1, operator skill raises the performance from a value of zero to a value of 0.01. Using the analysis in Section 4.1, the high level is equivalent to saving at least 0.016 ships per 1,000 miles of convoy duty, or to earning (strictly on a replacement basis) at least $96 per mile of convoy duty. Thus, while at first glance the improvement in this situation may appear to be insignificant, personnel selection and training, which achieved merely this increment, would pay for itself in fairly short order. For example, if the training, etc., cost $100,000 and it were given to three men in the crew, the cost would be recovered in one trip across the Atlantic Ocean.

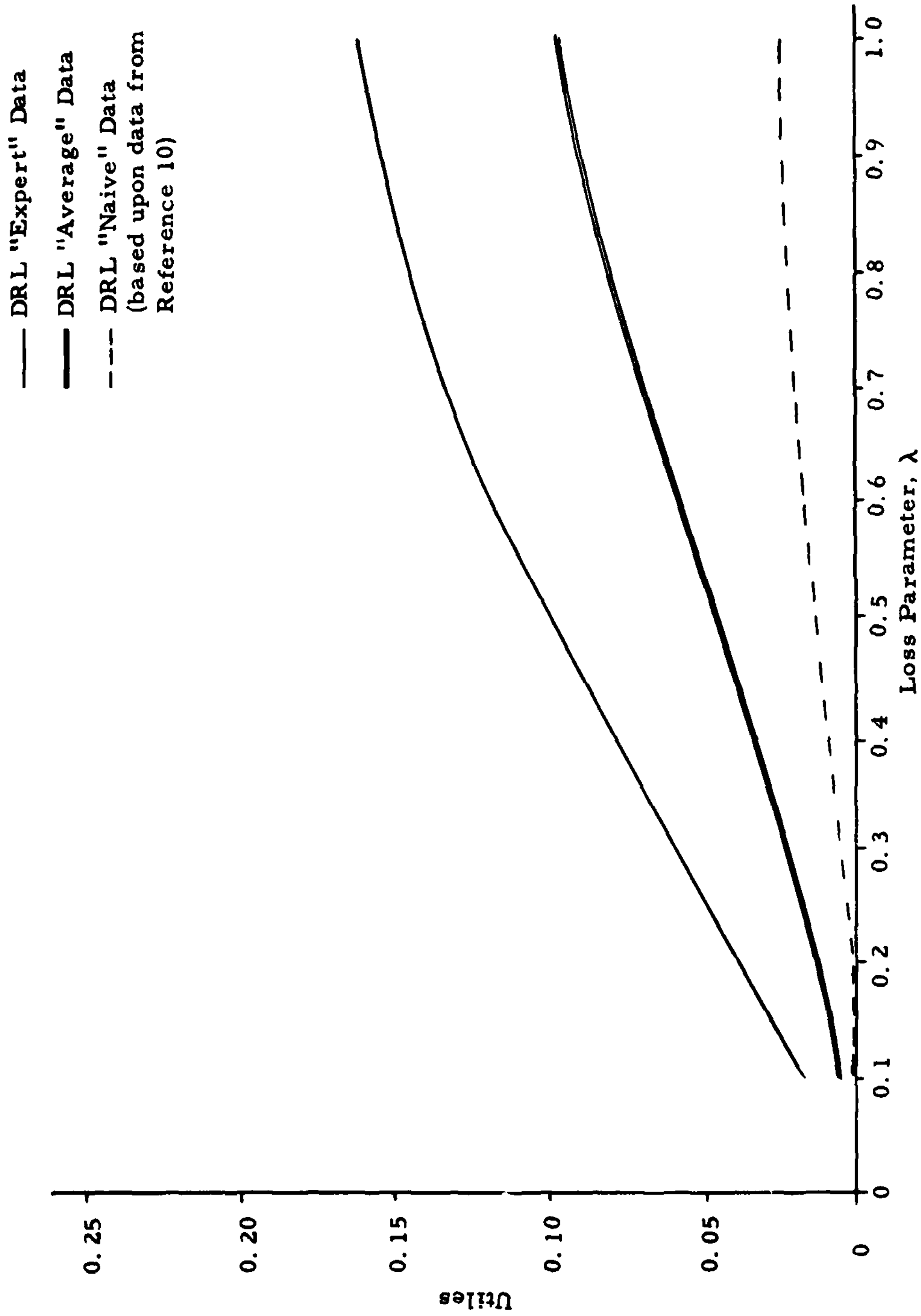

Figure 16. General values (V**) of System 6, using operators of three levels of skill. (C)

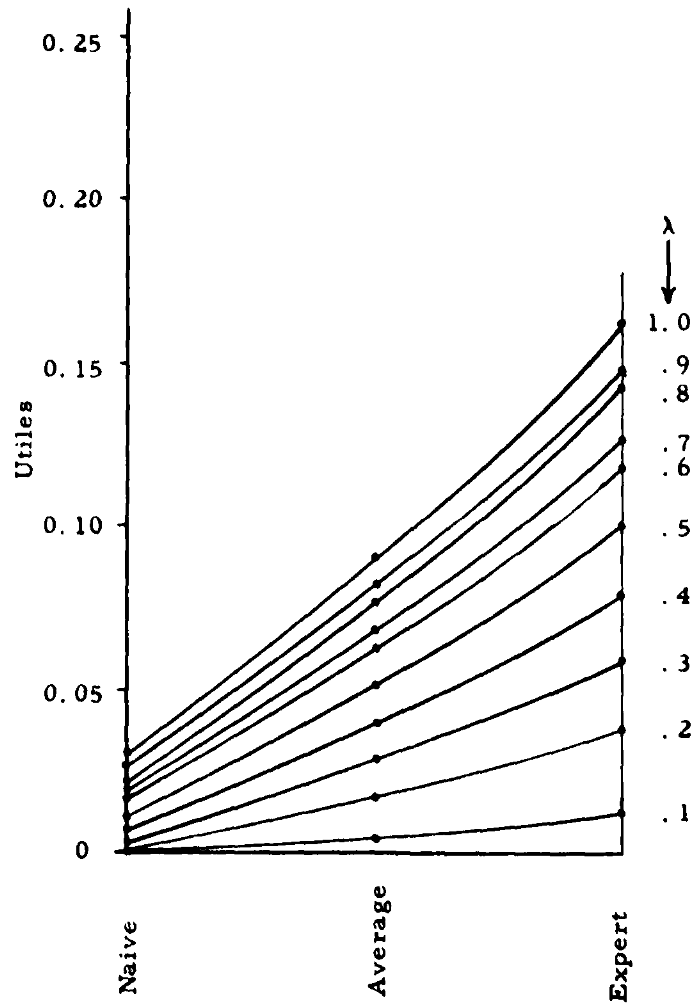

Figure 17. Relationships between general value and level of operator skill. (C)

4.4 Proposal for a Revised Classification Procedure (C)

Prior to the last section we have been considering classification systems in a fairly abstract sense. In particular, we have not specified the identity of the "decision maker": we have merely pointed out that the decision is made, wittingly or unwittingly, in accord with some decision rule with consequences which can be evaluated in terms of utility losses which may not be known by the decision maker.

The preceding section included explicit recognition of the fact that the sonar operator plays an integral role in associating received signals with clues; that is, in performing the first of the two functions associated with human implementation in Figure 1. In implementing this process, the sonar operator (or team of operators) does not act as a decision maker in the sense of selecting the terminal action. However, in the experiment in which the ROCs were generated, the man was not asked to associate observed signal characteristics with clues in the usual sense. Instead, he was asked to register one of ten possible integrated reactions.[10] Reactions number 5, 4, 3, 2, 1 were associated with increasing conviction that the contact was a non-submarine. Reactions 6, 7, 8, 9, 10 were associated with increasing probabilities that the contact was a submarine. The ROCs are determined from observations of the frequency of occurrence of each observation for each of the two basic conditions, sub and non-sub, known to the experimenter. This experimental procedure serves to isolate the human task of recognizing and categorizing received signal characteristics from the decision-making task of selecting a terminal action, thus providing the ROC, which is independent of the use of the decision rules and the associated problems of evaluating losses and the relative importance of data from sources other than the radar itself.

In the procedure used at sea, the information derived from the sonar is used as the input to a decision process in which the contact is classified as either a "possible sub" or as a "non-sub".[11] This process is usually referred to as carried out by "the sonar operator"; however, there is always a team of operators and in addition there are instances in which the captain, who acts subsequently as an independent decision maker, also participates in the initial decision task.[12] The decision rules used in selecting one of these two terminal actions are never completely formulated. Only some of the possible clue combinations are unequivocally associated with one of the possible actions. In addition, the clues themselves may be partially ignored or merely assimilated subjectively with no attempt at explicit clue identification.

Once the decision is made solely (or primarily) on the basis of the sonar-derived information, the captain has an opportunity to reverse or confirm the decision before taking additional action. In this second decision process, the captain has additional sources of information available; for example, other sensors such as radar, a plot of the target track derived from previous observations, and intelligence inputs.

This two-step process can be analyzed by noting that the captain treats the sonar operator's output as an observation. There are two such observations, an initial classification of the contact as a non-sub, which we designate $\tilde{S}$, and an initial classification of the contact as a sub, which we designate S. The captain has two terminal actions which we designate a and $\tilde{a}$ as before.

There are three significant decision rules:

d_1 specifies that $\tilde{a}$ is selected when the observation is $\tilde{S}$, and $\tilde{a}$ is selected when the observation is S (that is, if the operator says "S" the captain reverses the initial decision).

d_2 specified that $\tilde{a}$ is selected when the observation is $\tilde{S}$, and a is selected when the observation is S (that is, the captain confirms the initial decision).

d_3 specified that a is selected when the observation is $\tilde{S}$, and a is selected when the observation is S (that is, if the initial decision is $\tilde{S}$, the captain reverses it).

This procedure is demonstrably inferior to one which imitates the experimental procedure by allowing the sonar operator a variety of outputs, thereby providing the captain with a broader, more versatile set of decision rules.

Figure 18 illustrates the salient points in the argument. The experiments have demonstrated that the sonar operator is capable of generating a ROC curve such as that passing through the points (a) through (g) on the figure. The number of segments on such a ROC curve is determined by the number of output categories provided: in this illustration there would be seven. If the sonar operator must select either S or $\tilde{S}$, his behavior can be represented as the use of one of the decision rules associated with one of the points on the ROC. For illustrative purposes, suppose he selects the rule associated with point (e). The ROC being generated for the captain's consideration is the dotted line through (a), (e) and (g). In particular, d_1 is associated with (a), d_2 with (e), and d_3 with (g). If (e) is the point associated with the optimal rule and if the captain uses d_2 confirming the sonar operator's

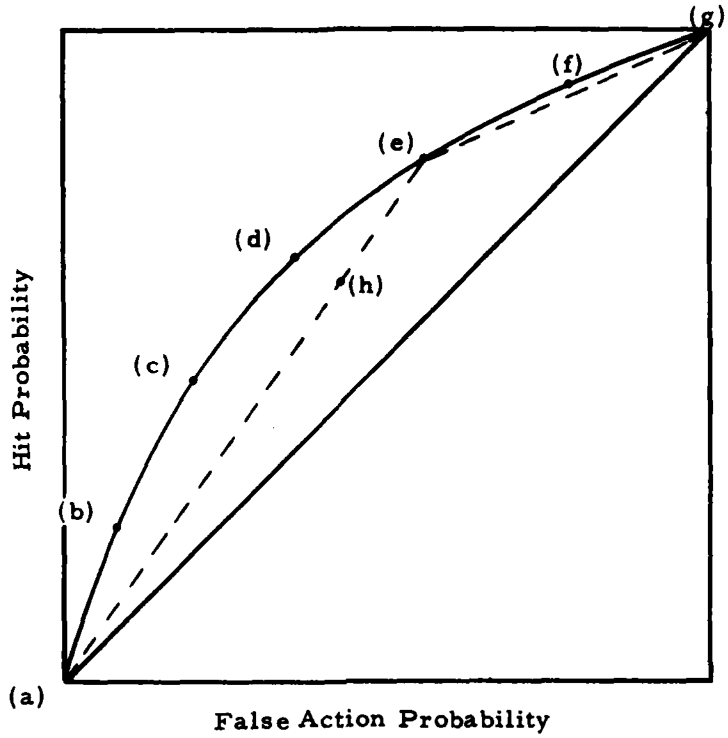

Figure 18. Effect of expanding sonar operator's output. (C)

choice, then the minimum risk is achieved but the captain's action is redundant. On the other hand, suppose that sonar operator has erred and (d) is actually optimal. The captain, despite his additional information, is still constrained to use either (e) or (a). At best if he is willing to use a mixed decision rule (and we have pointed out that this is very unlikely), he will use a rule associated with some point on the line segment between (a) and (e), say (h). However (h), or any other point on this line segment, is still inferior to (d) since all such points lie closer to the null system represented by the diagonal from (a) to (g). In other words, if (d) were optimal, (h) would be better than (e) but not as good as (d). This argument can be duplicated for any other initial choice and consequently the general conclusion is that a procedure in which the captain can use the full ROC is superior to the one he must use when the sonar operator's output is constrained to a choice between "possible sub" and "non-sub."

Additional insights can be obtained from Figure 19 in which the solid curve is a duplication of $V^*(p, \lambda)$, for $\lambda = 0.6$ from Figure 9. If the sonar operator selects S or S with no a priori knowledge of p, his best procedure is to use the maximin decision rule, d_6: this will limit the expected loss to at most V_1 on the figure. This then constrains the captain to three decision rules as discussed: in this case, his choices are d_1, d_6, and d_{10}. The resulting $V^*(p, \lambda)$ is the dashed line.

Since d_6 is optimal with no a priori knowledge, the captain cannot improve the situation except by using his additional information about p. Suppose the captain knows that $p < p_3$. With this information, he should be capable of improving the situation. For example, if he knows that $p < 0.2$, the captain should select d_3, ensuring a risk less than V_2 which, in turn, is significantly less than V_1. However, if the captain is constrained to work on the dotted curve, this a priori knowledge is of no use for his best available strategy is still d_6. In fact, from the figure, it is clear that the captain's optimal use of a priori information when the situation is the one described by the dotted curve is as follows:

1. If $p_4 < p \leq 1$, use d_{10}.
2. If $p_2 \leq p \leq p_4$, use d_6.
3. If $0 \leq p \leq p_2$, use d_1.

In other words, use of other information about the probability that the contact is a sub leads the captain to either completely ignore the sonar operator's classification or else to go along with it. This is highly inefficient compared to the situation represented by the solid curve when the sonar operator produces multiple outputs. For example, when $p_1 < p < p_3$, significant improvements can be obtained by using either d_5, d_3, or d_2 as appropriate, rather than either d_6 or d_1. Note that the situation cannot be improved by using a mixture of d_1 and d_6 since the point representing such a mixed decision must lie along a line such as CC' with a positive slope between the zero slope associated with d_6 and the slope associated with d_1. No point on such a line can be better than d_1 for Case (3) above, or better than d_6 for Case (2).

In summary, the captain uses his additional information about the nature of the contact most effectively when he is given access to the multiple outputs the sonar operator is capable of producing, rather than only the constrained duality, "possible sub" and "non-sub."

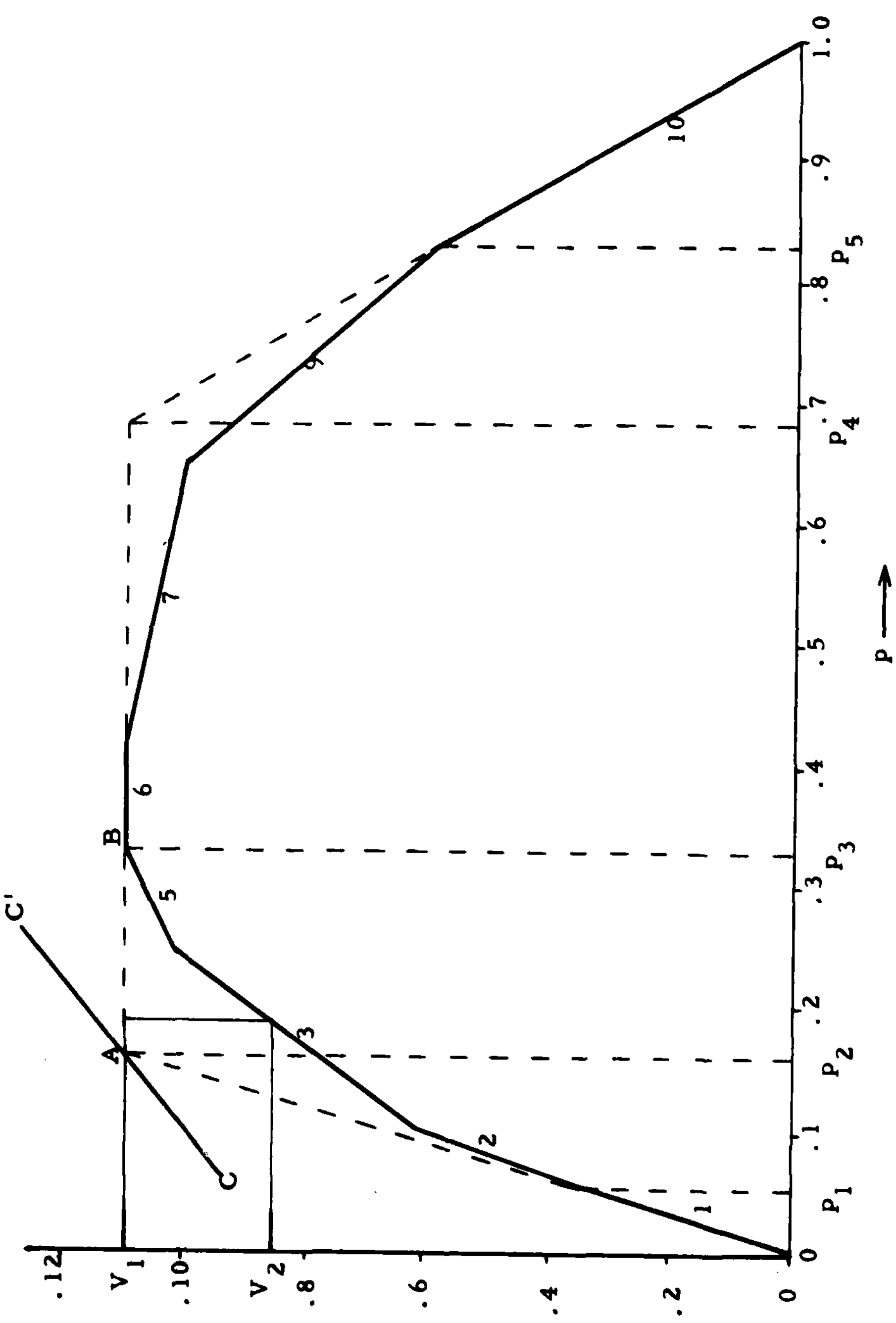

Figure 19. Risks with dual and multiple sonar operator outputs. (C)

4.5 Practical Procedures for the Captain (C)

In most circumstances the individual who selects the terminal action (that is, the "decision maker") is the captain of the ship. We have proposed that the output of the sonar operator (that is, the individual, or group, evaluating the information derived from the sonar signals) is a subjective estimate of the probability that the contact is a submarine (or a non-submarine). A practical way of implementing this procedure is to provide the sonar operator with a row of buttons and the convention that the left end of the row represents near certainty that the contact is a non-sub while the right end represents near certainty that the contact is a sub. The intermediate positions represent intermediated degrees of conviction, and the basic division between "sub" and "non-sub" is the middle of the row. For definiteness, we suppose there are ten such buttons, numbered from left to right (so number 10 represents near certainty that the target is a sub, etc.), and that each of these buttons switches on one of a row of lights observed by the captain when he wishes to ascertain the sonar operator's output. We also pointed out that the way in which the captain uses this information can be characterized theoretically as the selection of a decision rule and that there are optimum decision rules for given circumstances. We have not carried out a detailed investigation of the practical procedures which might be used to implement the process of selecting an optimal decision rule. However, the following general description of one possible procedure indicates that a practical methodology can certainly be developed.

A practical decision rule for the captain is:

> "Select action 'a' if light n, n+1, . . . ,10 is observed: otherwise select action 'a.'"

Initially the captain would select the value of n associated with the maximin decision rule. He would be taught how to select the proper n on the basis of practical, observable factors such as his ship's assignment (e.g., screen or SAU), the nature of the convoy (e.g., merchant ships or task force), and instructions from higher levels (who in turn might be considering observables reflecting the importance of the mission and the current exchange ratios).

As a second step, the captain would use a rationale which modifies the initial rule in accordance with more rapidly changing observables and personal judgment of their significance. For example:

> "If, on the basis of intelligence reports, you know you are now in a region in which encounters with submarines are likely to occur, shift the decision point down to n−1."

This corresponds to the theoretical concept of employing other information (or "a priori information") about the probability that the contact is a submarine. Another instruction might be

> "If the track looks like one which might be generated by a submarine preparing for an attack, shift the decision point down to n-1. If the attack looks imminent, shift it down to n-2."

This takes into account the effect described theoretically in terms of a decrease in λ corresponding to an increase in threat. It must be possible to compound such instructions; for example:

> "If you are in a region in which you are likely to contact submarines and the threat looks imminent, take action 'a' if any of lights n-3 through 10 are observed."

There will also be instructions which move the decision point in the other direction; for instance:

> "If ammunition is low (with 'low' appropriately defined in relation to the duration of the mission) shift the decision point from n to n+1. If ammunition is very low (again appropriately defined), shift to n+2."

This corresponds to adjusting to an increase in λ associated with a larger value of L_2.

A comprehensive set of instructions of this general kind and appropriate training in the interpretation and use of such directions will constitute a rational procedure which the captain can use to introduce his assessment of the situation and judgment into the selection of the terminal action.

There is at least one exception to this general procedure. When the captain's terminal action is a report to the OTC (or other higher echelon), his role is analogous to that of the sonar operator reporting to the captain. On the basis of the arguments already developed, it follows that the captain should not be constrained to a simple dichotomy (or even trichotomy) but should be allowed a fuller range of expression. The captain's output is then describable as an ROC presented to the OTC who is the selector of the terminal action. If the captain's action is significant, his ROC curve will lie further from the diagonal than that of the sonar operator and the value of the system (specific or general) will be higher when the captain's output replaces the sonar operator's. This is a subject which requires more detailed consideration beyond the scope of this report. However, our discussions of the sonar operator's functions can generally be applied directly to the role of the captain as an agent in the classification process rather than a terminal decision maker.

4.6 Research on Operator's Procedures (C)

The DRL data already cited demonstrates that the experimental determination of receiver operating characteristics will register differences between operators with different skills and classes of operators with different skills. Our analytic interpretation of these results has demonstrated that the value of the improvement obtained in going from "naive" performance to "expert" performance is large enough to justify extensive investment in training, personnel selection, and specialized equipment designed to upgrade classification performance by this amount. In other words, we have established that $V^{**}(\lambda)$ measures ASW classification system performance in terms of contribution to mission effectiveness and using this measure and the results of DRL's early experiments we have determined that sonar operators can have the skills required to make a significant contribution to overall ASW effectiveness. To capitalize on these results, we must now undertake the experimental program already proposed. In the light of the analysis presented above, we can now interpret some of the proposal's general statements about experimental objectives in much more precise terms.

The general objective of these experiments will be to identify the detailed relationships between human skills, attitudes, and other behavioral characteristics and performance in ASW classification as measured by V^{**} and the associated receiver operating characteristics. The repeatability, or more generally the stability, of the receiver operating characteristics is a particularly important consideration. Does V^{**} (and even more importantly, i^{**}) change unpredictably--as might happen if it is sensitive to mood, fatigue, and similar factors? If so, what are these factors? To what extent can the effects be mitigated by use of indoctrination, training, and personnel selection?

A second general experimental objective is to determine the relationships between the observed ROCs and the characteristics of individual operators. In particular it is necessary to confirm that the DRL observations of differences between operators of different skill are representative for larger samples of operators and for a larger variety of more realistic situations. For practical applications it is also important to determine whether or not the terminal decision rule should be modified to account for the characteristics of individual operators; that is, is i^{**} different for different operators? For personnel selection, it is important to determine if differences between operators depend upon human characteristics other than trainable skills.

(C) Another area to be investigated experimentally is the effect of various procedures upon the observed ROCs and associated system values. In particular, how useful (in terms of increasing system value) are the procedures associated with explicit identification of various clues (e.g., pip shape) and rules relating clue combinations to classification? Can operators using less explicit, more integrated subjective evaluations do better? How effective is the introduction of explicit aids such as HHIP? Would aids to remembering significant clue characteristics be more effective? The answers to these and similar questions will provide a practical basis for devising effective training procedures.

(C) The problem of setting up specific procedures for meeting these objectives is the next proposed step in this study. It is necessary to establish experimental plans utilizing specific equipment (including specific PME tapes if, as is most likely, a simulator is used) and suitably selected subjects.

REFERENCES

1. Wohl, J. G., et al., Anti-Submarine Warfare Training (U), Final Report, 14 September 1966, Prepared for: Chief of Naval Operations (OP 95D) Department of the Navy, Washington, D. C., Contract No. NObsr-91130.

2. Schumacher, G. P., *Summary Analysis of Clue Evaluators: HHIP and* MITEL Proceedings of Symposium on Active Sonar Classification (U), Volume I, Naval Postgraduate School, Monterey, California, 3,4,5 October 1967. Proceedings by Defense Research Laboratory, The University of Texas at Austin.

3. Green, David M., and Swets, John A., Signal Detection Theory and Psychophysics, John Wiley and Sons, Inc., New York, 1966, p. 34.

4. Green, David M., and Swets, John A., Signal Detection Theory and Psychophysics, John Wiley and Sons, Inc., New York, 1966, p. 88.

5. Luce, R. Duncan, and Raiffa, Howard, Games and Decisions: Introduction and Critical Survey, John Wiley and Sons, Inc., 1957, Chapter 2.

6. Raiffa, Howard, and Schlaifer, Robert, Applied Statistical Decision Theory, Harvard University: Graduate School of Business Administration, Division of Research, 1961, Chapter 4.

7. Blackwell, David and Grishick, M. A., Theory of Games and Statistical Decisions, John Wiley and Sons, Inc., 1954, Chapter 4.

8. Fishburn, Peter C., "Additive Utilities with Finite Sets: Applications in the Management Sciences," Naval Research Logistics Quarterly, Vol. 14, No. 1, March 1967, pp. 1-13.

9. McKinsey, J. J., Introduction to the Theory of Games, McGraw-Hill, 1952, p. 57.

10. Goodnow, R. K., A Study of ASW Sonar Operator Performance on a Realistic Classification Test: The Operator Classification Performance Test, Proceedings of Symposium on Active Sonar Classification, Volume II, Naval Postgraduate School, Monterey, California, 3,4,5 October 1967. Proceedings by Defense Research Laboratory, The University of Texas at Austin, pp. 41-69.

11. Office of the Chief of Naval Operations, Department of the Navy, Antisubmarine Classification Manual (U), NWIP 24-1A, pp. 1-5.

12. Sapp, Cmdr. E. W., Current Operational Classification Procedures. *Unpublished paper delivered at Symposium on Active Sonar Classification*, Naval Postgraduate School, Monterey, California, 3,4,5 October 1967.

13. Klipple, Alfred G. and Alvin J. Abrams, Accuracy and Consistency in Judging Active Sonar Classification Cues: IV. Graphic Cues. San Diego: U. S. Naval Personnel Research Activity, January 1966, pp. 27

14. Abrams, Alvin J. and Alfred G. Klipple, Accuracy and Consistency in Judging Active Sonar Classification Cues: III. Video Cue Judgments. San Diego: U. S. Naval Personnel Research Activity, pp. 35.

15. Accuracy and Consistency in Judging Active Sonar Classification Cues: II. Audio Cue Judgments, Technical Bulletin STB 66-14, December 1965, USNPRA.

16. Accuracy and Consistency in Judging Active Sonar Classification Cues: I. Overview of Problem and Summary of Findings, Technical Bulletin STB 66-12, November 1965, USNPRA.

DUNLAP *and* ASSOCIATES, INC.

DARIEN, CONNECTICUT 06820

One Parkland Drive
Area Code 203 655-3971
In N.Y.C., WYandotte 3-2464

Executive Offices

NEW YORK, NEW YORK 10017

200 Park Avenue
Area Code 212 661-2160

WASHINGTON, D. C. 20007

1050 Thirty-First Street, N.W.
Area Code 202 333-0100

BOSTON, MASSACHUSETTS 02116

607 Boylston Street
Area Code 617 262-6270

SANTA MONICA, CALIFORNIA 90404

1454 Cloverfield Boulevard
Area Code 213 393-0166

MANHATTAN, KANSAS 66502

200 Research Drive
Area Code 913 JEfferson 9-3565