BEHAVIORAL VALIDATION OF THE COMPUTER ASSISTED OPERATIONS RESEARCH FACILITY

DR. THOMAS J. HAMMELL
Eclectech Associates, Incorporated
and
MR. JOSEPH PUGLISI
Maritime Administration

INTRODUCTION

The Computer Assisted Operations Research Facility (CAORF) is a highly sophisticated ship maneuvering simulator owned and operated by the National Maritime Research Center. It was designed primarily as a high fidelity simulator to conduct a broad spectrum of applied maritime research. A secondary role has emerged for CAORF for the specialized training of VLCC masters in ports such as Valdez, Alaska. The cost of CAORF training for Valdez, in terms of resource conservation and other factors, has been demonstrated to be much less than the cost of comparable at-sea training. Recent interest by the maritime industry for the increased use of simulator training will require prototype development; CAORF is the ideal facility for this type of maritime training research. Hence, multiple roles have emerged for this facility: 1) to conduct research on operational maritime problems; and, 2) to develop maritime training programs and identify training device characteristics. Both of these roles exemplify aspects of the goal of resource conservation through simulation. The effectiveness of achieving this goal is addressed in this paper in terms of CAORF validation, the extent to which mate behavior on CAORF is similar to mate behavior at sea.

CAORF was designed as a research tool to enable the conduct of investigations that would be prohibitive to accomplish at sea. The extensive capabilities designed into CAORF for this purpose also enable its use as an advanced training device. A high fidelity wheelhouse environment is simulated, representing the state-of-the-art in merchant ship bridge simulation (see Figure 1). The major elements of simulation include: 1) the full compliment of ship instrumentation, such as radar and navigation equipment; 2) ship handling characteristics, such as wind effects, turning equations, and tugboat forces; and, 3) a comprehensive visual scene. The computer-generated visual imagery, subtending a 240° horizontal arc and a 24° vertical arc, is the most unique feature of CAORF. The full color visual scene can be constructed of any area and changed in real-time as a function of own ship and other ship actions. The visual scene can vary according to day/night, level of visibility, and color factors. It can include parts of own ship, other ships, coastline, bridges, buildings, oil rigs, and other visual objects.

Of equal importance to the quality of the simulation are the control and performance monitoring capabilities (see Figure 1). These are essential from both a research and training standpoint. The remotely located Control Station provides a complete set of indicators and displays for monitoring the actions of own ship and other ships. These include monitors of the visual scene, geographic plots, radar repeaters, and communication panels. Interrogation and control of all problem parameters may be accomplished from the Control Station. The Human Factors Station compliments the Control Station by providing a variety of remotely controlled audiovisual monitors for observing behavior on the bridge. These two stations permit remote observation and exercise control by a researcher or instructor. Furthermore, interaction with the mate or master on the bridge can be accomplished via CRT displays on the bridge and radio/telephone communication. Extensive computer facilities enable the on-line calculation of performance measures, provide for immediate feedback if desired, and provide for comprehensive post-exercise analysis. These extensive control, performance monitoring, and computational features, although designed for the conduct of research, are excellent training assistance capabilities. They provide the training program developer and instructor with a wide range of

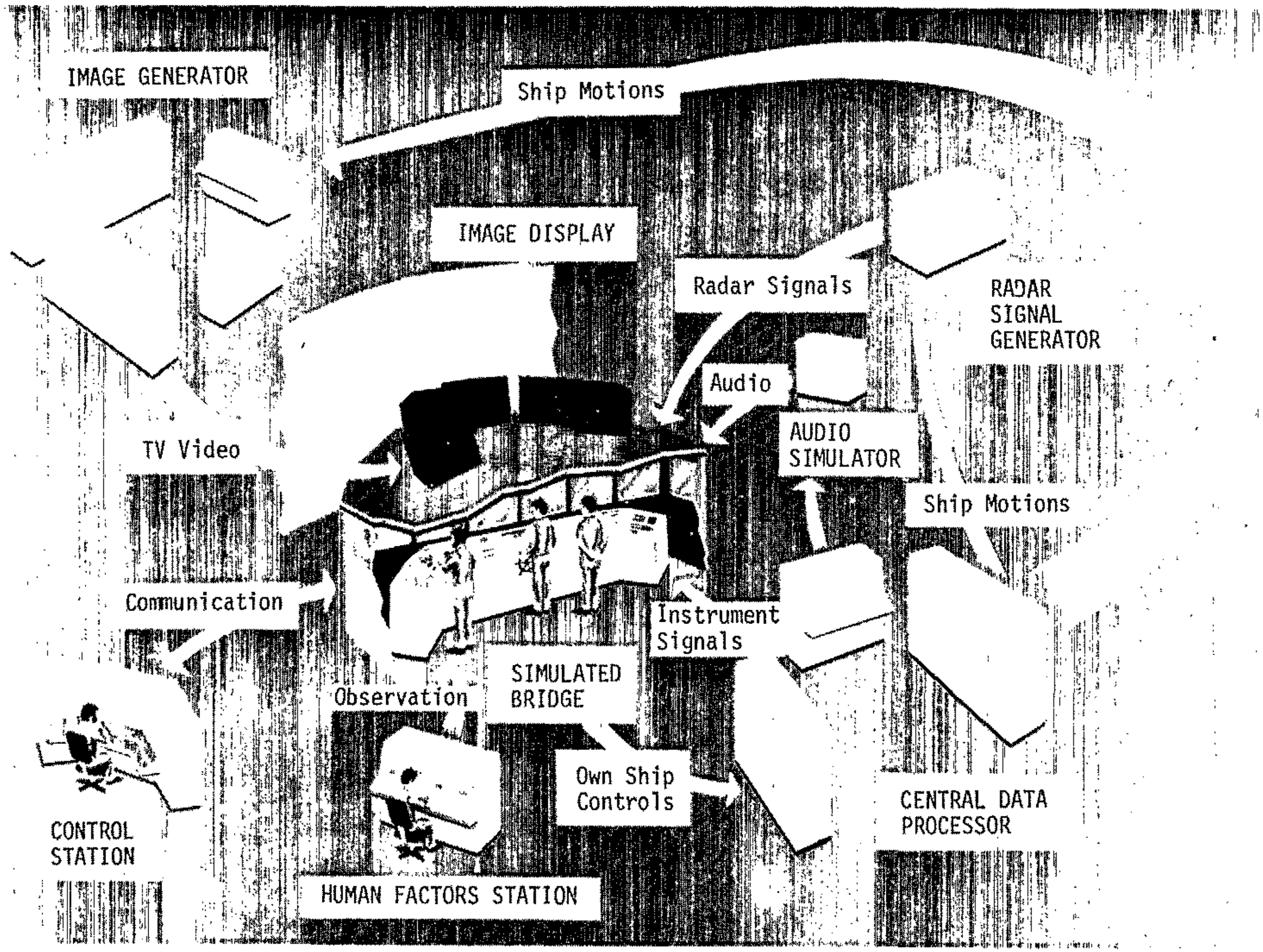

Figure 1. CAORF Subsystems

capabilities with which to develop and implement training.

The effectiveness of a high fidelity simulation facility in the conduct of applied research, or in providing advanced training, is dependent on the validity of elicited behavior. The purpose of incorporating a high degree of fidelity is to assure obtaining behavior on the simulator that is similar to behavior that occurs in the real world environment. Validation of the CAORF simulator was undertaken for this purpose, in essence to ensure that findings on CAORF have a positive relationship to those expected at sea.

The issue of simulator validation may be addressed in three distinct areas:

1. Hardware Validation - e.g., equations of ship motion.
2. Perceptual and Information Processing Validation - the mate's perceptual integration of simulated information (e.g., visual estimation of target range).
3. Behavioral Validation - the mate's behavioral functioning (e.g., the mate's radar-related behavior).

The third area of validation, mate behavior, is addressed by this paper. This area is of fundamental importance in that it may serve as a summary of the first two areas, as well as directly relate to ship performance.

This validation project, which was sponsored by the National Maritime Research Center, was tailored to investigate a variety of validation issues of concern to CAORF. The most important were:

1. The mate's level of activity and radar behavior as a function of situation conditions,
2. Examination of the existence and magnitude of a learning/experience factor when the mate initially reports on the CAORF bridge, and
3. Comparison of research findings between research performed at sea and on CAORF.

BEHAVIORAL VALIDATION APPROACH

Behavioral validation was accomplished by comparing the behavior of mates on CAORF with the behavior of mates in similar situations at sea (see Figure 2). An at-sea data base consisting of mate behavior was compiled from a variety of trips at sea. This computer data base also contains reconstructed target and own ship data for the various at-sea situations. Three four-hour watch scenarios were constructed for use on CAORF from these data. The watches were identical to situations that had occurred at sea near the Cheasapeake Bay, Matanilla Shoals in the Straits of Florida, and the English Channel. The behavior of mates who stood watch on the CAORF bridge during the constructed scenarios (ten mates total) was compared with the mate behavior observed during the similar situations at sea.

The three selected at-sea situations covered a broad range of conditions to provide for the general applicability of the validation results (see Figure 3). A variety of information was developed for each watch based on the at-sea situation: 1) own ship data; 2) target data; 3) environmental conditions; 4) communications guidelines; 5) geographical data; and, 6) ship information and materials. In essence, the mate reported for watch on the CAORF bridge is a duplicate of a watch that had occurred at sea.

The data collection methods employed on CAORF were identical to those used at sea. A trained observer recorded the pertinent information concerning the mate's tasks. A variety of task categories were identified for the at-sea data collection efforts, and also used to classify the tasks observed on CAORF. More than seventy distinct task categories were considered under the following major areas: 1) ship control; 2) navigation; 3) communication; 4) visual; 5) contact assessment; and, 6) logging and miscellaneous.

The validation analysis was conducted in three separate configurations:

1. CAORF Watch Behavior Versus The Identical At-Sea Watch Behavior- The advantage of this approach was that CAORF behavior could be compared with at-sea behavior under identical conditions. The disadvantage was that the behavior of only one mate per at-sea watch could be used for the analysis.
2. CAORF Situation Average Versus Similar At-Sea Situation Average - The three CAORF watches were partitioned into segments on the basis of the situation conditions; average behavioral data were calculated across the

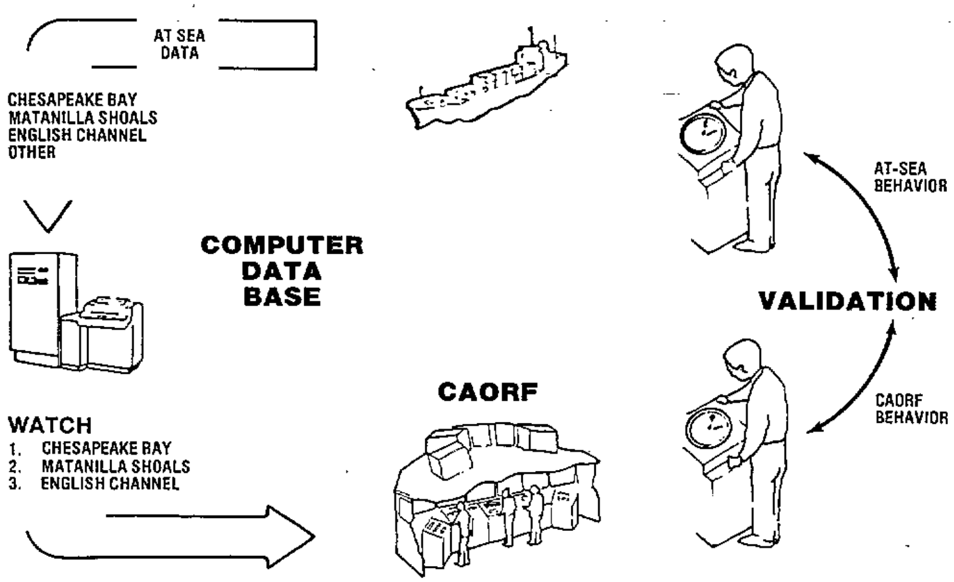

Figure 2. Behavioral Validation Approach

WATCH CONDITIONS		WATCH #1	WATCH #2	WATCH #3
LIGHTING	DAY		●	
	NIGHT	●		●
CONTACT LOAD	LOW	●	●	
	MEDIUM	●	●	●
	HIGH			●
NAVIGATION:	OPEN SEA	●	●	●
	COASTAL			
	RESTRICTED			
VISIBILITY:	UNLIMITED	●	●	
	LIMITED 3-7NM			
	LIMITED ≤ 3NM			●
COMMUNICATION:	LOW		●	●
	MEDIUM			
	HIGH	●		
RADAR NOISE:	LOW	●	●	
	HIGH			●
CURRENT SET:	LOW			
	MEDIUM	●	●	
	HIGH			●
NAVIGATION LOAD:	LOW		●	
	MEDIUM	●		
	HIGH			●

Figure 3. Watch Conditions

similar segments. The entire at-sea data base was likewise partitioned into segments, and behavioral averages developed. The advantage of this approach was that the at-sea data for each situation were based on the behavior of several mates.

3. Overall CAORF Average Versus Overall At-Sea Average - This approach developed average data across all CAORF watches, regardless of the situation conditions. The at-sea data were similarly averaged. This approach, although used to compare behavioral data, was primarily used for comparison of CAORF and at-sea ship performance.

The analysis investigated similarities and differences in mate behavior. Behavioral variables that received particular attention were: 1) Task Workload - the percentage of a mate's time that was spent doing particular tasks (e.g., a 20% radar workload means that the mate spent three minutes out of fifteen minutes performing radar tasks); 2) Task Frequency - the number of times in a given period that the mate performed particular tasks; and, 3) Task Duration - the average amount of time the mate spent performing a task. A contact load algorithm was used to provide an objective estimate of the loading placed on the mate by the distribution of ship contacts. The contact load algorithm was developed in an earlier study investigating standardized bridge design (Eclectech Associates, 1976). The contact load parameter is sensitive to the number of contacts (fixed or moving), their range, bearing, speed, and course. This algorithm was particularly useful in partitioning the watches into equivalent segments.

FINDINGS

Mate behavior on CAORF was found to be similar to mate behavior at sea. The level of mate activity, the type of tasks, and the manner in which the tasks were performed were found to change in accordance with the situation conditions, both on CAORF and at sea. A summary of the validation findings is presented in this paper. The interested reader is referred to Hammell (1976) for the complete report.

Overall comparisons between mate behavior on CAORF and at sea are shown in Figure 4. CAORF and at-sea workloads are averaged across all the data for three sets of situation conditions. The data are further broken down according to All Tasks and Radar Tasks categories. A relatively high correlation coefficient was obtained for these data (Spearman Rank Correlation Coefficient: T = 0.9, $p < .05$), demonstrating relative similarity in mate behavior on CAORF and at sea.

Behavior During the Watch

Mate behavior during the watch is exemplified by the data for Watch #2, Matanilla Shoals area, shown in Figure 5. The upper curve shows the dimensionless contact load as a function of time during the watch. A contact load of zero through ten indicates a low level; 15 through 30 indicates a medium level; 30 and above indicates a high level. The contact load reached a low-level peak around 0900 and a high-level peak around 1130. The lower curves show: 1) the mate workload on CAORF during the four-hour watch, for all tasks, and averaged over the six mates that stood this watch; and, 2) the mate workload at sea during the similar watch, for all tasks, and for the one mate that stood watch. Inspection shows that the mates usually had a higher level of activity on CAORF. Much of this difference may be attributed to a learning/experience effect, which is discussed later. The workload peaks near 0900 and 1115-1145 correspond to the contact load peaks; the workloads during these times are primarily due to contact assessment tasks (e.g., reflection plotting on radar). The workload peaks around 1000 were due to navigation tasks (e.g., plotting a LORAN fix). The similar workload trends of the CAORF mates and the at-sea mate during the watch (i.e., peaks around 0900, 1000, and 1115-1145) show the correspondence between CAORF and at-sea behavior as a function of the situation conditions.

Behavior Averaged Across Similar Situations

Greater confidence in the behavioral comparisons may be had when the CAORF data are compared with averaged at-sea data, reducing the potential bias. The three CAORF watches and the entire at-sea data base were partitioned into segments based on the situation

TASK	SEA CONDITION/ CONTACT LOAD	WORKLOAD (%)	
		CAORF	AT-SEA
ALL TASKS	OPEN SEA/LOW	27.0%	15.6%
	OPEN SEA/HIGH	48.2	32.60
	RESTRICTED WATERS/ LIMITED VISIBILITY	47.6	54.00
RADAR	OPEN SEA/LOW	12.0	6.00
	OPEN SEA/HIGH	34.4	20.00
	RESTRICTED WATERS/ LIMITED VISIBILITY	30.0	22.00

Figure 4. CAORF and At-Sea Workload Comparison Summary (Spearman Rank Correlation Coefficient: T= .9(p<.05).

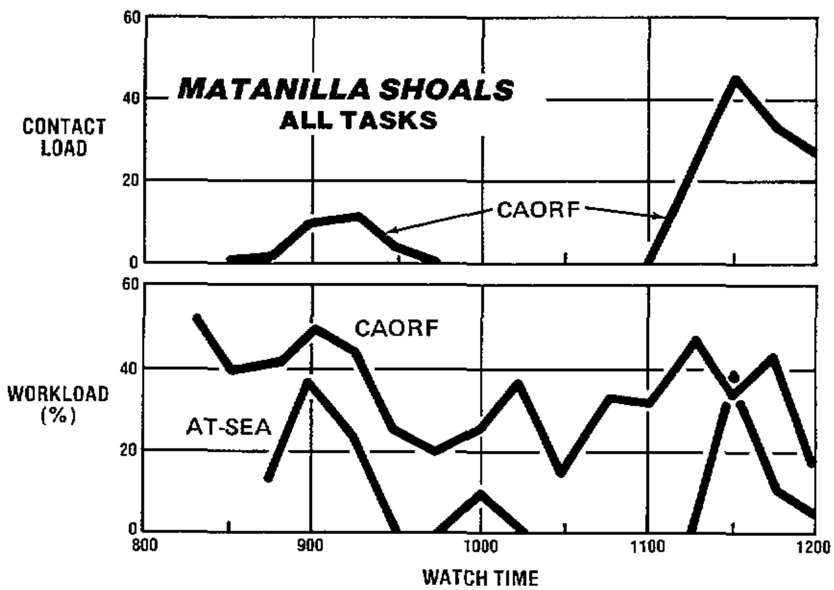

Figure 5. Contact Load and Mate Workload During Watch #2, Matanilla Shoals.

conditions. Three distinct groups of segments were identified from the CAORF watches on the basis of the situation conditions: 1) open sea, good visibility, low contact load; 2) open sea, good visibility, high contact load; and, 3) restricted waters, limited visibility, high contact load. The average workloads for all tasks during these segments show that the mate workload differences between the three groups of situations on CAORF are similar to those at sea (see Figure 6). These similarities extend to specific task areas as well (e.g., radar tasks).

Learning/Experience

Figure 6 shows that the workload in each group is greater on CAORF than at sea. This nearly constant offset is attributed to a learning/experience effect that occurs when the mate initially reports on the CAORF bridge. This effect is similar to what is expected to occur when a mate first reports on an unfamiliar ship. He has to learn the instrumentation, the location of the controls, and the setup of radar; he has to become familiar with the ship operating procedures, and so on.

The learning/experience effect has been hypothesized to occur in two parts - short-term and long-term. The short-term effect occurs at the beginning of his first watch, eventually settling down to a more normal level of activity. Figure 7 shows the mate workload with the short-term learning/experience period segmented. The three situation categories in this figure are: 1) learning/experience - the first hour of the CAORF watch, which always had a low contact load; 2) low contact load situations that occurred after the first hour; and, 3) high contact load situations. The relatively high workload that occurs during the short-term learning/experience period is evident from Figure 7. The CAORF mates are as active during this period as when under the high contact load conditions. The impact of this short-term learning/experiénce period is exemplified by the workload difference between this group and the low contact load group, both of which occurred under low contact load conditions. Removal of the short-term learning/experience effect has considerably reduced the difference between CAORF and at-sea workloads. The remaining difference is attributed primarily to the long-term learning/experience effect.

Comparison of the workloads for all tasks with those for radar tasks in Figure 7 reveals that radar behavior was not overemphasized by mates on CAORF. The higher CAORF radar workload is in proportion to the higher CAORF workload for all tasks. Hence, the radar workload received the same relative emphasis on CAORF as at sea. The absolute differences are attributed to the learning/experience effect on all tasks.

The behavior of one CAORF mate was investigated with regard to the long-term learning/experience effect. This mate stood the #2 Watch (Matanilla Shoals) twice, the two times separated by three months during which he had some additional experience on CAORF. Figure 8 shows this mate's radar workload during the first and second times he stood this watch (i.e., upper and lower curves, respectively). The radar workload for the mate at sea is also plotted. These curves show a substantial difference in the radar workloads between the mate's two watches, indicating a long-term learning/experience effect. Furthermore, the CAORF mate's second watch radar workload was quite similar to that which occurred at sea (i.e., almost no radar activity until the contact load increased near the end of the watch). These data should be cautiously interpreted since they were from only one mate on CAORF and one mate at sea. Nevertheless, they suggest a long-term learning/experience effect. Additional research should be conducted at sea to verify the learning/experience effects on mate behavior when he first reports on an unfamiliar ship.

Experimental Effects

The initial experiment conducted on CAORF investigated the benefits of electronic aiding on ship performance under a variety of situation conditions. Three levels of electronic aiding were studied: 1) visual only; 2) radar and visual; and, 3) collision avoidance system (CAS), radar, and visual. The CAS condition was found to result in superior performance based on a variety of measures.

One of the major purposes of this validation study was to investigate the

ALL TASKS

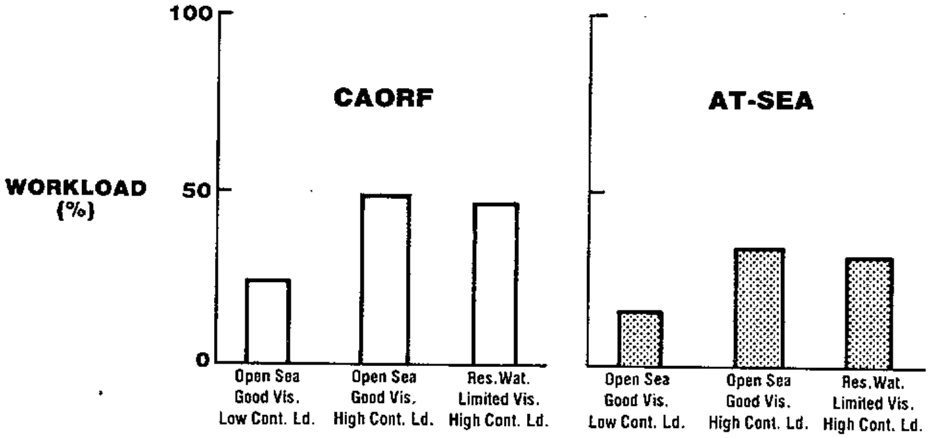

Figure 6. Average Workload for All Tasks, as a Function of Situation Conditions.

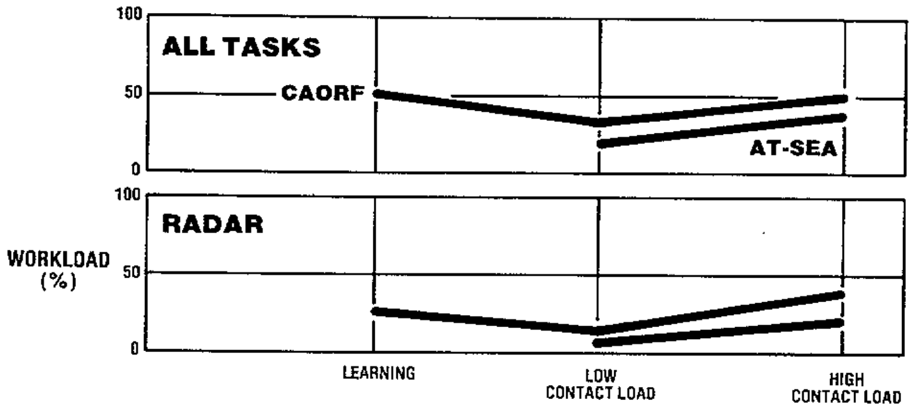

Figure 7. Average Workload as a Functioning of the Short-Term Learning/Experience Period, and Low and High Contact Load Situations.

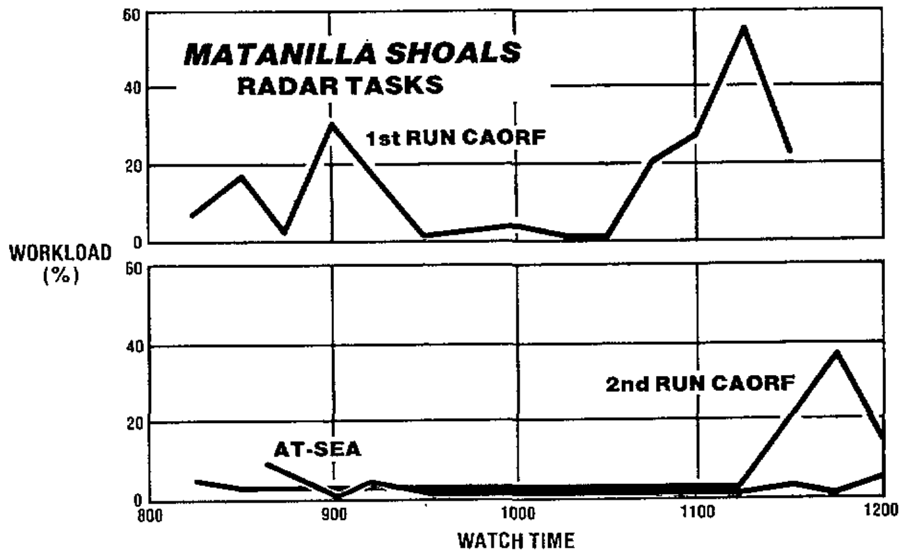

Figure 8. Mate Radar Workload During Watch #2, Matanilla Shoals, Showing the Long-Term Learning/Experience Effect.

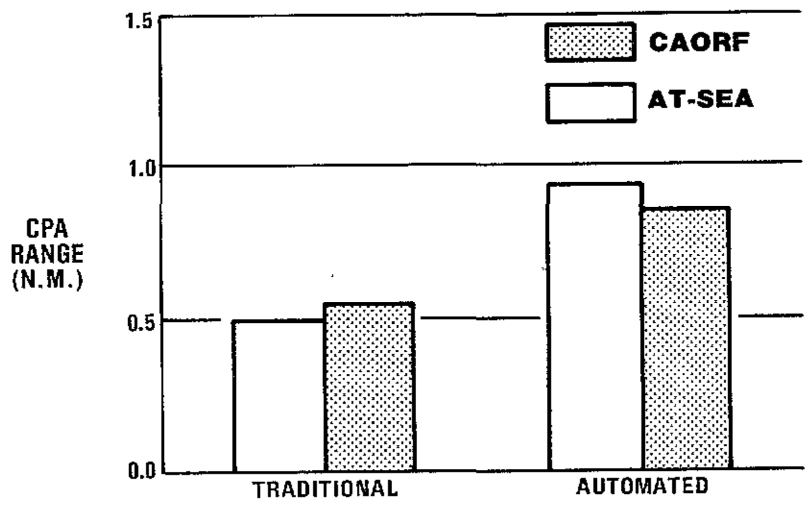

Figure 9. Range at Closest Point of Approach for Traditional (Radar) and Automated (CAS) Bridges, CAORF and At-Sea Data.

validity of CAORF experimental findings. This was accomplished by comparing experimental effects obtained from CAORF research with effects observed in the at-sea data. An important CAORF research finding supported by the at-sea data was that CAS equipped ships attained larger ranges at the Closest Point of Approach (CPA) than those ships equipped with radar only. Figure 9 shows the high degree of similarity in the ranges at CPA for CAORF and at-sea data for radar-only and CAS equipped ships. Of greatest importance, however, the at-sea data demonstrated an effect similar to that obtained from research on CAORF - that is, a greater range at CPA for the CAS equipped ships. The experimental effects obtained on CAORF, therefore, are substantiated by the effects observed in the at-sea data.

SUMMARY

The findings from this validation study demonstrate a degree of similarity between mate behavior on CAORF and mate behavior at sea. The mate's activity level on CAORF changed in accordance with situation conditions similar to behavior that occurred at sea. The CAORF mates were very active, at a level equivalent to that of the most difficult situations at sea. This generally high level of activity was attributed to a learning/experience effect that was predominant when the mates first reported on CAORF, and then subsided over time. This effect was similar to the expected behavior of mates at sea when first reporting on an unfamiliar ship. The mate's radar behavior on CAORF was also found to be similar to that occurring at sea, a function of the changing contact load conditions.

Experimental effects obtained in the initial CAORF experiment were verified by the at-sea data. These validation findings are, perhaps, the most important since they directly support the validity of CAORF research.

The CAORF simulator was designed primarily as a research tool. Its wide range of simulation, control, monitoring, and analysis capabilities make it an excellent training device as well. Although expensive, operation of CAORF is much less expensive than the operation of a ship at sea for research and training purposes, as has been demonstrated during experiments for the port of Valdez, Alaska. Other factors, such as the control of environmental conditions and the capability to train in failure modes of operation, weigh greatly in favor of the use of simulation. The acceptance of CAORF by the maritime community for research and training clearly demonstrates the need for resource conservation through simulation. The CAORF validation study represents a fundamental step in verifying the feasibility of simulation for solving a variety of problems, and hence, conserving resources.

The results of this validation study support the quality of the CAORF simulation and the resultant empirical research findings. This represents the first demonstration of the behavioral validity of a marine simulation facility.

REFERENCES

Eclectech Assoicates, Inc., "Standardized Bridge Program - Phase I - Ship Automation Benefits and Bridge Design Criteria," Office of Advanced Ship Operations, Research and Development, U.S. Maritime Administration, Department of Commerce, September, 1976.

Hammell, T. J., "Validation of Mate Behavior on the Computer Assisted Operations Research Facility (CAORF)" Eclectech Associates, Inc., Prepared for Grumman Data Systems Corportation, October, 1976.

ABOUT THE AUTHORS

DR. THOMAS J. HAMMEL is the Principal Scientist with Eclectech Associates, Incorporated. In this capacity, he is the Project Engineer on a submarine training system development effort; he also directs several training and research efforts being conducted for Chief Assisted Operations Research Facility (CAORF). He has directed a variety of projects concerned with the man-machine interface under the direction of the Navy, Maritime Administration, and other organizations. These include the validation of CAORF, the development of advanced training concepts and devices, the development of interactive computer models, and design of the man/computer interface. Prior to joining Eclectech Associates, Dr. Hammell was a Senior Research Psychologist with General Dynamics, Electric Boat Division. He received the B.E. degree in electrical mechanical engineering and M.M.S. degree in engineering psychology from Stevens Institute of Technology, and the Ph.D. in psychology from the University of Connecticut. He is a member of the American Psychological Association, Human Factors Society, and Military Operations Society.

MR. JOSEPH J. PUGLISI is a Project Manager with the National Maritime Research Center (NMRC), Maritime Administration. In this capacity, he manages projects in the areas of computer simulation, electronic navigation and communications at the NMRC Computer Aided Operations Research Facility. He joined the U.S. Merchant Marine Academy at Kings Point, New York in 1969 to work on the Savannah Nuclear Power Plant Simulator. When the NMRC was formed in 1971, Mr. Puglisi was assigned as Project Director on the Maritime Satellite NAV/COM System Experimental Program. His prior experience was as an engineer at the Naval Applied Science Laboratory in Brooklyn and the Naval Ship Research and Development Laboratory at Annapolis. Mr. Puglisi received a B.S. degree in electrical engineering from City College of New York and the M.S. degree from New York University and is presently in a Ph.D. program in computer science at Polytechnic Institute of New York.

PAPERS PUBLISHED BUT NOT PRESENTED