AD-A265 457

Report No. CG-D-07-93

MODELING TECHNIQUES FOR SHIPBOARD MANNING: A REVIEW AND PLAN FOR DEVELOPMENT

John D. Lee
and
Thomas F. Sanquist

Battelle Human Affairs Research Centers
Seattle, Washington

FINAL REPORT
FEBRUARY 1993

This document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161

Prepared for:

U.S. Coast Guard
Research and Development Center
1082 Shennecossett Road
Groton, Connecticut 06340-6096

and

U.S. Department Of Transportation
United States Coast Guard
Office of Engineering, Logistics, and Development
Washington, DC 20593-0001

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

This report was directed and sponsored by the Coast Guard Research and Development Center. This report does not constitute a standard, specification or regulation.

SAMUEL F. POWEL, III
Technical Director
United States Coast Guard
Research & Development Center
1082 Shennecossett Road
Groton, CT 06340-6096

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
CG-D-07-93		

4. Title and Subtitle	5. Report Date / 6. Performing Organization Code
Modeling Techniques for Shipboard Manning: A Review and Plan for Development	5. Report Date: February 1993 6. Performing Organization Code:
7. Author(s): John D. Lee and Thomas F. Sanquist	8. Performing Organization Report No.: R&DC 04/93
9. Performing Organization Name and Address: Battelle Human Affairs Research Centers P.O. Box C-5395 4000 NE 41st Street Seattle, Washington 98105-5428	10. Work Unit No. (TRAIS) 11. Contract or Grant No. 13. Type of Report and Period Covered: Final Report
12. Sponsoring Agency Name and Address: Department of Transportation U.S. Coast Guard Office of Engineering, Logistics, and Development Washington, D.C. 20593-0001 U.S. Coast Guard Research and Development Center 1082 Shennecossett Road Groton, CT 06340-6096	14. Sponsoring Agency Code

15. Supplementary Notes

The Coast Guard technical contact and COTR is Dr. Anita Rothblum, (203) 441-2847.

16. Abstract

The worldwide trend in shipboard manning is toward increasingly smaller crew sizes. One concern regarding smaller crews is to what extent they may compromise ship safety and the ability to respond in an emergency. In order to address this concern, one needs a systematic method of establishing safe crew complements that can be applied on an international basis. Manning models can provide a flexible, economic method for establishing safe manning levels.

This report critically analyzes existing manning models, both within and outside the maritime domain. A detailed examination of six models developed for the maritime domain illustrates how limits associated with tools and content, as well as the complexity and diversity of issues involved with manning requirements, make it difficult to use any of these six models as a technical basis for all ship manning decisions. Rather than develop a single, comprehensive manning model, it is suggested that the wide variety of issues related to safe manning may require a number of specialized models, each of which is used to examine specific issues.

A plan for manning model development for the U.S. Coast Guard is presented. A relatively simple shipboard work-hours model is suggested, which can be expanded and enhanced to consider other issues, such as maintenance or emergency procedures. Criteria for model development are explained, as are procedures for model validation.

17. Key Words	18. Distribution Statement
manning crew complement model validation	Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. SECURITY CLASSIF. (of this page)	21. No. of Pages	22. Price
UNCLASSIFIED	UNCLASSIFIED		

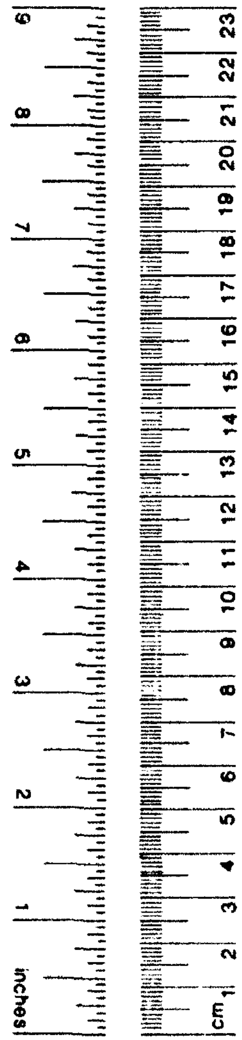
Form DOT F 1700.7 (8/72) Reproduction of form and completed page is authorized

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (WEIGHT)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (EXACT)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in = 2.54 (exactly)



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (WEIGHT)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	0.125	cups	c
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (EXACT)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F

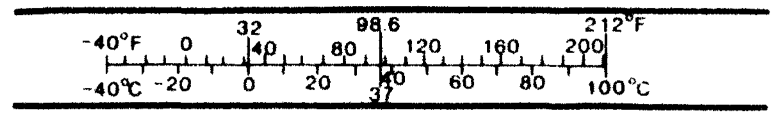


TABLE OF CONTENTS

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of the following individuals: Dr. Marc Mandler, Dr. Anita Rothblum, CAPT William Gibbs, Dr. Alexander Landsburg, Ms. Mireille Raby, Dr. Walter Pearson, Ms. Susan Czarny, and Mr. Richard Ecker. In addition, we would like to express our appreciation for the cooperation and support of CDR Myles Boothe, Mr. Chris Young, LCDR Marc Wilson, and the other members of the Human Factors Coordinating Committee of the United States Coast Guard, Washington, D.C.

EXECUTIVE SUMMARY

RATIONALE AND PURPOSE

The issue of shipboard manning scales is a complicated legal, economic and human factors engineering question. The worldwide trend in shipboard manning is toward increasingly smaller crew sizes - in the case of some Japanese ships, crews are composed of as few as 11 persons. Crew sizes of this level are made possible by advanced technologies permitting unmanned engine rooms and one-man bridge operation, as well as reductions of deck crew for cargo handling. A fundamental question that arises as a result of such crew reductions is the extent to which smaller crews compromise ship safety and the ability to respond in an emergency. This question has not been addressed satisfactorily. There is a need for a systematic method of establishing safe crew levels that can be applied on an international basis. A number of authors share this opinion (Seitz, 1981; Knudsen and Mathiesen, 1987; Froese, 1987; Joseph, 1987; Gaffney, 1987), but at this time no comprehensive approach has been established.

Several issues make the need for a model of ship manning increasingly important. Most generally, they involve understanding how reduced or modified staffing influences ship safety. The general issue of ship safety involves several more specific concerns, such as the effects of increased technology, emergency response effectiveness, and maintenance capabilities. In establishing safe crew levels, the government and industry need to consider demands on the crew, each vessel's technology, type of service, crew skills, and quality of management and management programs. Evaluating how these, and other issues affect ship safety provides a crucial step toward developing a technical basis for Coast Guard policy.

In evaluating ship safety, a variety of techniques can be used including: simulator experiments, sea trials, accident analyses, and manning models. Manning models circumvent some of the problems involved with other approaches. In many cases manning models provide a more flexible, more economic, and less dangerous method for establishing safe manning levels.

Once manning models have been created, experiments can be conducted to show how changes in the manning structure might affect ship safety. Models combine data to produce estimates of workhour requirements, timelines of personnel activities, estimates of mental and physical workload, and estimates of accident probabilities. These predictions can then be evaluated as measures of effectiveness, providing estimates of ship safety. With these predictions, users of the manning models can evaluate how changes in crew configurations, manning levels, and voyage profiles affect ship safety. Because of the need to collect data and validate the output of a manning model, models would be used in conjunction with other techniques, such as sea trials and simulator experiments.

Although manning models offer a very promising approach to developing a technical basis for ship manning, any model developed will have limitations. These limitations include the types of scenarios it accommodates, decision support capabilities, ease of use, and ability to consider execution of shipboard tasks over time. This report presents a framework for comparing different modeling approaches so that accurate models can be developed to address issues concerning safe ship manning.

MODELS IN MARITIME AND NON-MARITIME DOMAINS

The history of models used to investigate the role of humans in domains other than the maritime industry illustrates the utility of models in assessing the safety and efficiency of complex systems. While the specific purposes and domain characteristics differ from the manning issues of concern to the Coast Guard, the generality of the approaches and the similarity of some characteristics of these domains make some of these approaches applicable to the maritime industry. More generally, the success of many of these models illustrates the feasibility of developing models to establish a technical basis for manning decisions. In addition, the limitations of these models illustrate some of the potential problems that a manning model may face. In general, the limitations of these models fall into two categories: (1) limitations of the modeling tools and, (2) limitations of the content incorporated into the model. If either the tool or content fails to match the characteristics of the system to be modeled, then the manning requirements produced by the model may seriously underestimate the actual manning requirements. A review of models developed for non-maritime applications illustrates the great diversity of possible models, and the need to develop models with capabilities tuned to the specific issues involved with ship manning.

A detailed examination of six models developed for the maritime domain illustrates how limits associated with tools and content, as well as the complexity and diversity of issues involved with manning requirements, make it difficult to use any of these six models as a technical basis for all ship manning decisions. However, these models were designed for specific functions and in most cases they were successful, illustrating the potential of developing a technical basis for ship manning. Therefore, this review shows that it may not be possible to develop a <u>comprehensive</u> manning model. Instead, the wide variety of issues related to safe manning may require a number of <u>specialized</u> models. While no one model is comprehensive enough to examine all issues regarding manning decisions, many of these models offer the potential to resolve specific issues, such as the ability of a manning structure to meet workhour requirements. Taken together, they illustrate the feasibility of developing manning models to estimate safe manning requirements.

A PLAN FOR MANNING MODEL DEVELOPMENT

Developing a comprehensive approach for estimating manning requirements must go beyond previous studies. Much of this effort involves determining what factors need to be included in the model and then collecting data to model these factors. An economical approach to this challenge is to develop a relatively simple model that determines manning requirements based on workhour limits, then enhance and expand this model as factors are identified and data are collected. The efforts of Argent Marine and the National Research Council (NRC) indicate that a model based on workhour limitations is practical, and collecting the requisite data is feasible. Likewise, the studies by Kurke (1961) and Madni, Chu, Purcell, and Brenner (1984) show that detailed models, developed to address specific issues, are also feasible.

The development of manning models must meet several criteria. First, a general functional analysis needs to be performed. This analysis would specify the general functions that the ship personnel must fulfill, providing the basis for estimating ship manning requirements. A general functional analysis would provide a template for estimating the time, expertise and number of crew members needed to fulfill the duties and tasks identified in the functional analysis. Second, a profile of factors needs to be compiled that enumerates what should be considered when the Coast Guard evaluates proposed manning structures. This profile defines the model content, and should include guidelines regarding model scope, level of detail, and organization, as well as requirements for addressing such issues as: the limits of human performance, effects of automation, capability to respond to emergencies, ability to accommodate unplanned maintenance activity, and capacity to comply with statutory provisions and international standards. Third, developing a validation process will ensure that the recommendations provided by the model accurately reflect the conditions onboard ships, ensuring safe manning levels. This validation process should provide a structured method that can apply to the analysis of all aspects of model development and use, including: data collection, modeling tool selection, model content specification, and model use and interpretation. It would provide a method for the Coast Guard to evaluate data and models that industry might submit in support of manning proposals. This validation process would extend to sea trials and evaluation of safety records after the introduction of changes to manning complements. The fourth criterion requires development of a baseline estimate of manning requirements for a wide variety of ships, based on the functional analysis of ship requirements. This estimate would depend on approximations of task times, but would provide a starting point for more refined estimates. Any successful approach to a technical basis for determining minimum manning requirements should begin by meeting these four criteria.

MODEL VALIDATION

A crucial step in the development of any of these models is validation. The validity of a model depends on its ability to replicate the behavior of the actual system. Because any model represents a simplification or abstraction of the actual system, the behavior of the model will never be isomorphic with the behavior of the actual system. Therefore, any model will never achieve perfect validity; however, a model's behavior must replicate the behavior of the actual system closely enough to support decision making. Thus, while validity of a model will never be complete, it should be measured with respect to the purpose of the model and the decisions it supports. Model validity can be defined and measured in a number of ways. This report presents several definitions of validity from experimental psychology, along with validation methods derived from systems engineering. A specific example of some of the steps included in the validation of a functional model to evaluate workhour requirements include: 1) verify functional listing, 2) verify capabilities of modeling tool, 3) compare output with validated measures of effectiveness, 4) verify time and skill expertise estimates associated with shipboard functions against standards and against actual practice during sea trials, 5) check software configuration to ensure accurate estimation of workhour requirements, 6) evaluate the match between the stated purpose of the model and the current application, and 7) establish whether special restrictions should be placed on the conclusions drawn from the model.

CONCLUSIONS

This report addresses the need to develop a technical basis for ship manning. The goal of such a technical basis is to provide information that decision makers need to balance the effects of technological advances and economic pressures with ship safety. The models presented in this report illustrate the potential for estimating the effects of manning structures on ship safety. Although a variety of methods could serve this purpose (such as adjusting ship manning and evaluating the rate of subsequent accident reports), manning models offer a safer, more efficient, and more cost-effective approach. Therefore, the Coast Guard should begin to develop manning models.

The development of manning models should begin with the construction of a functional model to evaluate workhour requirements. Constructing such a model, based on the data and techniques of the NRC and Argent Marine, provides a cost-effective way of developing a basic manning model and evaluating the need for additional data collection and model development.
Developing a functional model to evaluate workhour requirements would provide a foundation, upon which the Coast Guard could build more detailed models. This basic model could then be elaborated upon with more specific models that indicate the ability of a proposed crew to accommodate factors such as emergencies, maintenance, automation, and port operations. The additional models would elaborate on the basic workhour requirements model by expanding its scope and verifying its assumptions. The report includes an outline describing how several models might be constructed, the first providing the foundation, upon which subsequent models can build. The proposed models include: (1) the ability of ship personnel to meet workhour requirements specified by regulations, (2) the ability of ship personnel to manage emergency situations, (3) the effect of crew reductions, facilitated by automation, on the navigation performance, and (4) the ability of a reduced crew to meet demands of scheduled and unplanned maintenance.

Thus, the development of a technical basis for ship manning should begin by defining detailed specifications of a manning model to estimate workhour requirements, based on a functional analysis of the ship system. These specifications, developed in close coordination with the users of a manning model at Coast Guard headquarters, would identify its scope, flexibility, operation, and role in the regulatory process. The most practical method of generating these specifications would be through an iterative process, where manning model prototypes are developed and presented to the Coast Guard staff. These prototypes could be quickly refined, based on Coast Guard critiques, and presented for further evaluation. The end product of this process would be a basic manning model that provides a first step towards a technical basis for manning levels. In addition, it will focus subsequent data collection and model development.

I. INTRODUCTION

1.1 BACKGROUND

The issue of shipboard manning scales is a complicated legal, economic and human factors engineering question. The worldwide trend in shipboard manning is toward increasingly smaller crew sizes - in the case of some Japanese ships, crews are composed of as few as 11 persons. Crew sizes of this level are made possible by advanced technologies permitting unmanned engine rooms and one-man bridge operation, as well as reductions of deck crew for cargo handling. A fundamental question that arises as a result of such crew reductions is the extent to which smaller crews compromise ship safety and the ability to respond in an emergency. This question has not been addressed satisfactorily. There is a need for a systematic method of establishing safe crew levels that can be applied on an international basis. A number of authors share this opinion (Seitz, 1981; Knudsen and Mathiesen, 1987; Froese, 1987; Joseph, 1987; Gaffney, 1987), but at this time no comprehensive approach has been established.

Although a variety of methods can be used to evaluate manning requirements (such as adjusting ship manning and evaluating the rate of subsequent accident reports), the 1990 National Research Council (NRC) report illustrates the limitations of these techniques. For example, the many concurrent changes that might accompany manning reductions make it very difficult to identify a causal link between manning reductions and subsequent changes in accident rates. One promising method of addressing these problems is through the development of a model[1] that predicts the consequences of manning changes on system[2] safety.

Mechanical, electrical, material and structural engineers have used models to evaluate the consequences of design modifications on ship performance for many years. Most recently, advances in computer aided design, particularly finite element analysis, have led to more optimal designs, which have greatly increased ship efficiency. As technological and economic pressures combine to modify manning structure, more advanced models of the human element aboard ships may provide a method to evaluate the safety consequences of proposed changes.

---

[1] In the context of this report, "model" refers to a representation of a system that translates input variables, such as changes in manning, into output variables, such as changes in ship safety. In essence, a model provides a simplified, abstract, representation of an actual system, for the purpose of evaluating system effectiveness.

[2] In the context of this report, "system" refers to a collection of interacting components and may be limited to bridge personnel interacting with navigation aids, all ship personnel interacting with ship hardware, or even the interaction between ship personnel, ship hardware, and shoreside personnel and hardware. Therefore, a "system" can include a variety of factors, and can be used to illustrate many points where the exact definition of the "system" is not important.

The 1990 NRC study entitled "Crew size and maritime safety" states:

> In establishing safe crew levels, the government and industry need to consider demands on the crew: each vessel's technology, type of service, crew skills, and quality of management and management programs. Systems engineering methods, including functional task analysis, offer an objective basis for such determinations. The model developed is one approach to implementing systems engineering approaches to determining safe ship manning. (NRC, 1990; pp. 73)

As this quote from the NRC report states, a variety of techniques, encompassed by systems engineering, might address a wide variety of factors to establish safe manning levels. Using functional task analysis to develop a manning model of the ship system, the NRC implemented *one* approach to estimating manning levels, based on task time aggregations. The NRC could have chosen from many other systems engineering approaches to produce a variety of models, each with different capabilities and limitations. Any model developed will have limitations that include the types of scenarios it accommodates, decision support capabilities, ease of use, and ability to consider execution of shipboard tasks over time. This report presents a framework for comparing different modeling approaches to determine which modeling techniques might best address the issues concerning safe ship manning. The complexity of a ship system makes it impossible for any one model to portray the whole system accurately. It would be very difficult to construct a single that would incorporate all the factors concerned with ship manning. Therefore, the basic assumption of this framework is the need to match modeling capabilities to the issues that need to be resolved. In this light, the worth of a model lies in the ability to represent crucial characteristics of the system under consideration, rather than simply the number of features it includes.

1.2 SCOPE OF REPORT

The scope of this report includes a discussion of capabilities and limitations of previous manning models, and provides a framework for evaluating and developing future maritime manning models. We do not provide a fully developed manning model. Instead, we describe approaches and assumptions that must be considered in model development. The report focuses on four major areas. First, this report provides a short summary of modeling efforts in other industries such as military systems, nuclear power plant operations, process control plants, and the aircraft industry. This review illustrates the diversity of modeling techniques and model content, showing that no one model will answer all manning questions. Second, because no one model answers all manning questions, this report describes a framework for evaluating the validity of potential manning models. It describes the capabilities and limitations of modeling tools, as well as the consequences of including various kinds of information in a model. Third, this framework is used to evaluate six existing models developed for the maritime industry, illustrating the capabilities and limitations of these models. Fourth, the report describes four specific potential manning models, together with a description of an overall strategy to guide the development of a series of manning models.

1.3 THE NEED FOR A MANNING MODEL

Several issues make the need for a model of ship manning increasingly important. Most generally, they involve understanding how reduced or modified staffing influences ship safety. This general issue involves several more specific concerns, such as the effects of increased technology, emergency response effectiveness, and maintenance capabilities. Developing an understanding of how these, and other issues affect ship safety provides a crucial step towards developing a technical basis for Coast Guard policy.

Introducing increasingly sophisticated technology poses two problems. First, technology may change the nature of the work, shifting the role of people from direct involvement with the ship to monitoring the performance of automation and intervening when it fails. The consequences for human effectiveness in this new role are poorly understood (Sheridan and Hennessy, 1984; Moray, Ferrel, and Rouse, 1990). Second, automation promises to reduce the physical and mental workload of ship personnel; however, the capability of automation to reduce workload in a way that promotes ship safety, especially in abnormal situations is poorly understood. Research in the aviation and process control domains suggests that introducing automation often compromises system safety rather than enhancing it (Woods, Potter, Johannesen, and Holloway, 1991; Wiener, 1985). In many instances designers automate what is most easily automated, leading to a patchwork of systems that inhibit rather than support personnel. In addition, automation often operates smoothly during normal operations, but fails when the personnel need it most, during abnormal operations, leading to workload peaks greater than those experienced without automation (Wiener, 1989). Similarly, Bainbridge (1987) suggests that highly automated systems may lead to a loss of skills needed when the automation fails. Therefore, simple assumptions of workload reductions accompanying the implementation of automation may not provide a valid estimate of the manning requirements needed to ensure ship safety. This implies that a predictive model of the human-machine system may promote the successful introduction of automation, and better estimates of manning requirements. Sheridan (1987) developed the concept of supervisory control to describe how humans interact with semi-automated systems. Sheridan and Hennessy (1984) discussed modeling approaches appropriate to this situation and concluded that, while human interaction with highly automated systems lacks a comprehensive model capable of making exact predictions, specific models have been used successfully in the process control, aviation, and manufacturing domains (Sheridan and Hennessy, 1984).

In addition to the problems associated with introducing new technology, emergency response capabilities represent an aspect of ship operations that may be especially sensitive to changes in manning. Emergency response requires a coordinated effort of the crew, often in unfamiliar and adverse situations. This response may require more people than normal operations do, and the demands on these people may exceed those of routine activities. In addition, the ability of personnel to react efficiently to emergencies has immediate implications for the safety of the crew and for the integrity of the ship. Therefore, a model that could relate the consequences of modified manning structures to the ability of the ship's personnel to mitigate emergency situations would provide crucial information on how changes in manning structure would affect overall ship safety.

Like emergency response capabilities, the ability of ships' crew to maintain the ship effectively may be particularly sensitive to changes in the manning structure. Because the ship's safety is directly dependent on the quality of the maintenance, conditions which lead to inadequate maintenance require identification. Preventive and shore-based maintenance both act to alleviate unexpected maintenance demands. However, the effectiveness of these techniques in conjunction with reduced manning has not received a detailed examination. A model incorporating the interaction of shore-based and preventive maintenance with modified manning structures, will reveal the capabilities of different crew structures, combined with different maintenance schedules, to maintain a ship effectively. For manning models to provide a technical basis for manning decisions, they must address the impact of increasing technology, emergency response capabilities, and maintenance effectiveness, as well as other specific issues that relate to the more general question of how changes in manning structure affect ship safety.

A wide variety of techniques, ranging from accident analysis to simulator experiments and manning models, can be used to evaluate ship manning. One of the simplest methods is sea trials, where a proposed crew structure is placed on ships and the resulting changes are carefully monitored. These changes might include increased workhours, increased workload (as measured by subjective questionnaires), decreased safety (as reported by onboard observers), and increased statutory violations. These changes could be used to determine if the changes in ship manning have compromised ship safety. While sea trials might provide valuable data regarding the effect of changes in the manning structure, this technique should not be used exclusively because the consequences (collisions and groundings) of inadequate manning are so severe.

Because of the inherent dangers of sea trials, other methods of evaluating manning proposals need to be considered. For example, critical incident analysis (semi-structured interviews focusing on dangerous or unusual situations) could provide a better understanding of the current demands placed upon ship personnel. Data collected through critical incident analysis would include a catalog of potential problems caused by inadequate manning. For example, critical incident analysis might reveal particular tasks or situations that require an unusually large number of people. While critical incident analysis provides useful guidance regarding ship manning, it alone does not include a method of summarizing the data to provide precise guidance for manning decisions. Simulator experiments are another alternative that could help determine manning requirements. More specifically, they could help specify how many people are required on the bridge to manage the ship under a variety of conditions. Experiments can be designed to evaluate precisely the effects new technology on the workload and performance of different configurations of bridge personnel. Unfortunately simulator experiments are limited to evaluating bridge personnel, providing little information on the potential effects of different manning structures on the workload of other ship personnel.

Computer-based manning models can provide another alternative for evaluating ship manning. Manning models represent simplified versions of the actual ship system. Therefore, once manning models have been created, experiments can be conducted to show how changes in the manning structure might affect ship safety. Model development requires data collected from task analyses, expert judgment, historical records, simulator experiments, and operating procedures. Models combine these data to produce estimates of workhour requirements, timelines of personnel activities, and estimates of mental and physical workload, estimates of accident probabilities. These predictions can then be evaluated as measures of effectiveness, providing estimates of ship safety. With these predictions the users of the manning models can evaluate how changes in crew configurations, manning levels, and voyage profiles affect ship safety. Because of the need to collect data and validate the output of a manning model, models would be used in conjunction with other techniques, such as sea trials and simulator experiments.

As a supplement to other methods used to verify ship safety, models could provide information needed to resolve issues regarding manning levels. The following list of potential questions and the information that manning models may provide serve to illustrate the capabilities of manning models:

1) Is the proposed manning structure able to attend to ship functions within the statutory workhour limitations?

Given estimates of the time and skill requirements associated with ship functions, a manning model could identify instances where the time and skill demands cannot be met by the proposed manning structure.

2) How would the incapacitation of one or more crew members affect the ability of the remaining crew to satisfy ship functions within the workhour limitations?

Given estimates of the time and skill requirements associated with ship functions, a manning model would provide a timeline of the hours worked each day by each crew member for the duration of the proposed voyage. Timelines could be generated for a variety of combinations of personnel incapacitation, showing how robust the manning structure is to disruption.

3) Would the proposed crew be able to accommodate emergency situations?

Based on emergency procedures followed by the crew, combined with time and expertise estimates of the tasks involved, a manning model would show how changes in the manning structure affect the interaction and coordination of the activities of all ship personnel involved in responding to an emergency. These effects could be combined with model output that shows changes in emergency response times to determine whether the proposed manning structure can adequately respond to emergencies.

4) Changes in manning structures require that certain maintenance activities occur during port calls. However, some unplanned maintenance may need to be done during voyages because of safety considerations. Is the proposed manning structure able to attend to unplanned maintenance activities?

Given the probabilities of various maintenance tasks, their time demands, their skill requirements, and the proposed manning structure, a manning model could calculate the probability that a failure would occur to which the proposed crew could not respond. Given that the crew could respond, the manning model could also calculate the expected delay associated with the repair.

5) Does increased automation on the bridge justify manning reductions?

Given a detailed cognitive task analysis of the navigation, management, and communication tasks, a manning model would provide data concerning workload and the ability to manage the ship. A simple estimate of mental workload based on the time required divided by the time available would indicate possible operator overload. In addition, increased time to respond to potential collision situations would also reflect the safety consequences of reducing bridge manning in response to increased automation.

The answers to each question are not directly tied to ship safety. Instead, the users of the model must interpret the results and decide whether the changes predicted by the model compromise ship safety. In some instances this may be relatively simple. For instance, given the current regulations, a manning structure that leads the model to predict 15 hour workdays may be deemed unsafe. In other instances interpretation of the model output is less clear. For instance, increased automation may lower mental workload, resulting in either increased safety or decreased safety. While overload can lead to unsafe situation, underload can also compromise safety. Very low workload could actually threaten ship safety by increasing boredom and compromising vigilance. Other measures, such as the probability that ship personnel will be able to respond to unplanned maintenance, are more straightforward. However, it is not clear what probability represents an unacceptable threat to ship safety. An initial task in developing a technical basis for manning decisions would be to agree upon measures of effectiveness that would combine the output of manning models and provide criteria for judging "ship safety".

1.4 THE ROLE OF MANNING MODELS IN THE REGULATORY PROCESS

A manning model may play a variety of roles in the Coast Guard's regulation of ship manning. At one extreme the Coast Guard could collect data, construct a model, maintain a model, and develop regulations based on the model output. At the other extreme, the Coast Guard could leave it to the maritime industry to collect data, construct models, maintain the models, and develop manning levels based on the output of these models. In this situation the Coast Guard would evaluate models developed by industry to ensure that their proposed manning levels are accurate. In the first case, the Coast Guard acts to **develop** and **operate** manning models, and in the other they would act to **evaluate** and **guide** the maritime industry in developing their own manning models. Each of these scenarios has benefits and limitations and requires different types of commitments from the Coast Guard.

The scenario where the Coast Guard develops and operates a manning model would require a substantial investment in the development of a manning model. Since this model must apply to a wide variety of ships, in a wide variety of circumstances, the model must provide a framework that would be able to represent any ship that might apply to the Coast Guard for a reduction or modification in their crew structure. As part of developing and operating a manning model the Coast Guard might also need to collect data. Alternately, the Coast Guard might require ship owners to provide data. As the 1990 NRC study suggests:

> In the vessel certification process for vessels employing new manning concepts, each operator should conduct a thorough assessment of shipboard functions and tasks required by the particular vessel and should submit a functional analysis (with specified crew numbers and structure, skills and training, voyage profiles, and operational and maintenance plans) to the Coast Guard for approval. (NRC, 1990; pp. 87)

Using these data as inputs into a manning model, the Coast Guard could provide a technical basis for manning level certification. This scenario has several requirements. First, the Coast Guard must develop a manning model, or series of manning models, capable of representing a wide variety of ships and operating circumstances. This model (or models) must be able to incorporate a large profile of factors that might influence safe manning. For example, the model should be flexible enough to include factors such as: human performance limitations, special qualifications, unplanned maintenance, emergencies, and statutory requirements. Second, the Coast Guard must either collect data regarding task time and expertise requirements, or provide industry with a generic template for this data collection that would then be "filled in" and submitted to the Coast Guard for approval. Third, after developing a new manning level the Coast Guard must evaluate whether the workload estimates produced by the model are accurate. This validation process might be guided by a standardized procedure or checklist, and could involve direct observations during sea trials, or examinations of vessels' safety records. Part of this validation process might include verification that the crew allocation and task times on an actual ship match those included in the model. The results of this validation procedure would indicate whether the manning level is appropriate and could signal the need to revise the manning model.

In the other scenario, the Coast Guard would not base manning decisions on a model developed and operated by the Coast Guard. Instead, they would guide the maritime industry to develop the appropriate manning models that would be submitted to them when they seek approval for a new manning structure. The Coast Guard would then evaluate the validity of the manning model as part of the certification process. This scenario would require the Coast Guard to develop a prototype manning model (or models) to be disseminated to industry. Companies could then use this prototype as a guide for developing models that they would submit to the Coast Guard for approval. This scenario goes a step beyond the NRC recommendation that suggests industry might submit a functional analysis in support of their manning proposal, since it requires them to develop and operate a complete manning model. Therefore, the Coast Guard would not have to develop and operate a manning model, but they would need to guide industry and evaluate their models and the model output. Although this scenario alleviates the Coast Guard of much of the work involved with developing manning models it has several requirements. First, the Coast Guard must develop explicit guidance concerning what types of models should be developed. As part of this guidance the Coast Guard should provide model templates or prototypes that include a generic identification of tasks and functions that should be considered. In addition they should provide a repertoire of modeling tools and guidance for selecting the most appropriate tool for model development. Second, the Coast Guard should describe a general profile of factors that need to be included in manning models. This profile might include the effects of human performance limitations, special qualifications, unplanned maintenance, emergencies, and statutory requirements. Third, the Coast Guard needs to develop a validation process to evaluate the models submitted by companies. This validation process should examine the modeling technique adopted, the underlying assumptions, the data collection methods, and the output of the model. As a final step in this process, a standardized procedure or checklist could guide direct observations during sea trials, or examinations of vessels' safety records, with the results supporting decisions concerning manning levels and model validity.

Whether the Coast Guard acts to **develop and operate** manning models, or to **evaluate and guide** the maritime industry in developing their own manning models, several criteria must be met. First, a general functional analysis needs to be performed. This analysis would specify the general functions that the ship personnel must fulfill, providing the basis for estimating ship manning requirements. A general functional analysis would provide a template for estimating the time, expertise and number of crew members needed to fulfill the duties and tasks identified in the functional analysis. Second, a profile of factors needs to be compiled that enumerates the factors that should be considered when the Coast Guard evaluates proposed manning structures. This profile defines the model content, and should include guidelines regarding model scope, level of detail, and organization, as well as requirements for addressing such issues as: the limits of human performance, effects of automation, capability to respond to emergencies, ability to accommodate unplanned maintenance activity, and capacity to comply with statutory provisions and international standards. Third, developing a validation process will ensure that the recommendations provided by the model accurately reflect the conditions onboard ships, ensuring safe manning levels. This validation process should provide a structured method that can apply to the analysis of all aspects of model development and use, including: data collection, modeling tool selection, model content specification, and model use and interpretation. It would provide a method for the Coast Guard to evaluate data and models that industry might submit in support of manning proposals. This validation process would extend to sea trials and evaluation of safety records after the introduction of changes to manning complements. The fourth criterion requires development of a baseline estimate of manning requirements for a wide variety of ships, based on the functional analysis of ship requirements. This estimate would depend on approximations of task times, but would provide a starting point for more refined estimates. The remainder of this report examines the feasibility of developing a manning model (or models), and provides suggestions to guide model development so that these four criteria can be met.

This report provides a basis for selecting the most appropriate modeling tools, and criteria for specifying model content. In addition, it outlines a strategy for model development, and presents validation methods that can verify the quality of the final model and its recommendations. Chapter II reviews modeling efforts in other domains, and shows how the use of models in other domains might guide the use of models in the maritime industry. Chapter III presents a brief summary of classifications of modeling tools and their content, illustrating how various model characteristics influence model validity. Chapter IV reviews current maritime models, illustrating how future manning models might build upon these efforts. Chapter V illustrates how modeling tools combine with content to form models capable of addressing issues central to ship manning. This section also provides a strategy for model development and validation.

II. APPLICATION OF MANNING MODELS IN NON-MARITIME DOMAINS

Human factors modeling techniques have been used for many years, not only in the maritime industry, but also in the aviation, nuclear, military and process control industries. A review by Dutton and Starbuck (1971) found 2024 studies, published before 1971, that discussed modeling of humans either as individuals or within systems. The availability of computing resources has increased the rate of model development since 1971 (Elkind, Card, Hochberg, & Huey, 1990). Because of the large number of models, this report includes only a brief summary, examining several representative models, rather than providing a comprehensive catalog of models. More comprehensive reviews include Pew and Baron (1983); Pew, Baron, Feehrer, and Miller (1977); Booher and Hewitt (1990); Farooq and Dominick (1988); Woods, Roth, and Hanes (1986); Bogner, Kibbe, Laine, and Hewitt (1990) and Elkind, Card, Hochberg,, and Huey (1990). While most of these models do not address the exact issues of the maritime industry, the different techniques and approaches provide a base on which the maritime industry can build. In addition, the success of many of these models illustrates the feasibility of developing models for maritime applications.

2.1 COMMERCIAL AIRCRAFT

Similar to the issues involved with ship manning, the aircraft industry has faced requests for reduced manning in conjunction with increasingly automated systems. During verification of the flight deck crew reduction (from three to two), models were developed to analyze the effect of crew reduction on aircraft safety (Straus and Stein, 1981). This research has continued as more automation finds its way into the cockpit, endeavoring to predict pilot workload as a function of a variety of factors including manning level, automation, and flight scenarios. One of the primary criticisms of the original research was the tendency to amalgamate workload estimates across time, making it impossible to evaluate the potential for short-term occurrences of extreme levels of workload (Straus and Stein, 1981). In this case the evaluation criterion was an estimate of the average workload experienced by the pilots. Ideally, the output would consist of a timeline that identifies and estimates peak workload levels. More recent efforts have addressed this issue (Parks and Boucek, 1989). In addition, validation tests during the development of the Boeing 757 and 767 airplane programs showed the timeline analysis technique predicted flight data with more than 90% accuracy. Furthermore, estimates of pilot workload, gathered through subjective scales, physiological measures, and secondary tasks, had correlation coefficients with the predicted workload that ranged from 0.51 to 0.91.

Although the original aircraft manning models depended on a timeline approach (Parks, 1978), other efforts include task network models (Laughery, Drews, and Archer, 1986) and optimal control theory approaches (Pew and Baron, 1983). (See Appendix B for a detailed description of these approaches.) Specifically, Baron et al. (1980) developed PROCRU (Procedure-Oriented Crew Model), a highly sophisticated model that supports systematic investigation of issues concerning how design changes of commercial aircraft influence performance and safety during the landing phase of flight. This highly complex model not only represents the motor responses of pilots, but also the cognitive activity involved in goal selection and procedure execution. The data requirements of this model are extensive. First, PROCRU requires a system model that represents the dynamics of the plane. Second, it requires a representation of the cognitive and decision making aspects of human performance, including decision time and processing effort. Third, the model must represent communication between crew members. Fourth, the model must include crew members' knowledge of the state of the world, conditions of the system, and goals. The output of PROCRU includes a wide variety of performance metrics that include a complete description of the simulated path of the plane. It also provides timelines of crew activity, showing how certain segments of the flight might overload a given crew member. This output allows designers to perform a sensitivity analysis to determine how design parameters affect operator workload and system performance. The sophistication of the output of PROCRU enables it to support designers in their development of cockpit instrumentation and control algorithms for aircraft.

PROCRU represents one of the most complex models of human-machine interaction; however, as with other comprehensive engineering-based models, it requires a more complete validation. In other words, the predictions of PROCRU may not correspond to actual human behavior. Optimal Control Theory, upon which PROCRU is based, has been validated in a large number or contexts. In addition, many of the submodels included in the model have been validated and used elsewhere. However, PROCRU itself has not been completely validated. While a number of the submodels and assumptions that form the basis of PROCRU have been validated, this does not guarantee the validity of the overall model. The severe technological and economic problems associated with collecting data required to validate a model like PROCRU make it unlikely that it will ever receive a complete validation (Baron and Corker, 1989).

In addition to PROCRU, Human Operator Simulator (HOS) predicts human performance through a complex set of micromodels that can predict the effects of human performance limitations on system effectiveness. The micromodels include representations of short term memory, control manipulation, eye movements, and decision time. The data for the micromodels come from empirical studies, while a detailed task analysis specifies how the micromodels will interact. The output of this model includes predictions of system performance and potential for human errors. Lane et al. (1981) identified several applications where HOS has supplied valuable data concerning different design alternatives. HOS has been used to simulate a variety of commercial and military systems, not only guiding designers' choices, but validating the model. The most fruitful application may be in detailed examinations of specific design alterations, and not the effects of modifications to a manning structure. While the specific aim of HOS and PROCRU may differ from those of manning models in the maritime domain, the success of these models illustrates the feasibility of developing formal models of human behavior. This suggests that the behavior of ship personnel can also be modeled to evaluate manning proposals.

2.2 NUCLEAR POWER PLANTS

Several modeling efforts have addressed the effect of human behavior on the performance of nuclear power plants, but no models have been used tc evaluate staffing requirements directly. One of the most popular attempts to estimate the influence of human reliability on the performance of nuclear power plants was THERP (Technique for Human Error Rate Prediction) (Swain and Guttmann, 1983). The objective of THERP is to incorporate the likelihood of human errors into the engineering estimates of plant failure. To do this THERP requires a detailed task analysis that specifies operator actions. Data regarding the potential for errors for many of these tasks has been collected primarily from subjective estimates. Although this database has not been substantiated, it enables THERP to combine data regarding error probabilities for individual tasks in event trees that illustrate the probability of various system states, given the probability of human error. Since THERP considers only manual control and low level perceptual errors, rather than decision making errors, it may underestimate the actual influence of human reliability on plant performance. Even with these limitations, THERP has been used to estimate the affect of human reliability on system effectiveness in a variety of settings, most importantly the nuclear industry. In an experimental examination of THERP, Kirwan (1988) compared the predictions of THERP to actual data and found a strong correspondence, illustrating that THERP has the potential to model failure rate of operators' procedural activities.

Because analysis of nuclear power plant accidents has revealed that decision making failures have been an important cause of plant failure, compared to failures in procedural activities, researchers have tried to model the cognitive functions involved in human decision making (Woods, Roth, and Pople, 1987). The Cognitive Environment Simulation (CES) provides a model of human error in the area of intention formation. It predicts errors based on a mismatch between cognitive resources and task demands. Therefore, this model includes a representation of the physics that guide the evolution of plant failures, and the operator's knowledge and skills that might apply to mitigating the consequences of these failures. The output of this model consists of simulated decision making behavior. By varying the knowledge base used in the simulation, the model can reveal the types of errors humans with different level of expertise might commit. In addition, this model provides estimates of decision times that reflect process dynamics and expertise. This model represents an extremely high level of modeling sophistication that would likely exceed the requirements of most manning decisions. As such, it demonstrates the potential for developing models that address complex decision making processes, if they are identified as being important in manning decisions. This work represents the state-of-the-art in modeling human decision making. As such, it is a complicated model and its use requires experts familiar with simulation. While the model could not be directly applied to the maritime domain, it demonstrates the possibility of using models to estimate the reliability of human decision making. Future versions of this model will become increasingly sophisticated as the number of incidents covered expands and a wider variety of factors that influence decision making, such as reliance on automation, are included in the model.

Overall, techniques like THERP have been widely used to estimate the effect of humans on system reliability and could easily be applied to the maritime domain. However, these techniques apply only to well-defined sequences of activities or procedures. When activities are not well defined, and cognitive activities of the operators must be incorporated into the analysis, other techniques must be used. CES illustrates an example of how human decision making activity might be included in a reliability analysis. However, this model has not been thoroughly validated or extensively used. Additionally, the complexity of this approach makes development and validation of this type of model difficult and expensive. Therefore, even though CES is one of the few techniques able to evaluate decision making reliability it may not be cost-effective to extend it to the maritime domain.

On the more specific issue of manning levels, the nuclear industry has sponsored very little research. Current manning regulations focus on the control room, neglecting other areas of the plant. In addition, regulations concerning the control room specify manning levels that depend on simple distinctions between plants, such as the number of reactors and the number of control rooms. Currently, the nuclear industry routinely staffs plants with more personnel than specified by NRC regulations. This indicates that the regulations do not accurately reflect the demands on the operators. At the same time there is no guarantee that industry has accurately determined adequate staffing levels. The lack of a systematic technique to evaluate manning requirements is not limited to the U.S. nuclear industry. Melber et al. (1983) reviewed international personnel and staffing regulations and found a similar situation. No technical basis has been developed, and manning requirements generally depend on: (1) prior experience with comparable fossil fuel plants, (2) empirical reviews of plant performance, and (3) union demands and labor laws. No systematic method exists to translate task demands into manning requirements. Thus, the nuclear industry has developed sophisticated technologies for evaluating human reliability for specific procedures, but no systematic technique has been used to estimate manning requirements.

2.3 MILITARY COMMAND AND CONTROL

The military has developed a number of manning and workload evaluation models that illustrate how automation influences the performance of command and control systems. Command and control systems oversee the planning and execution of tactical activities associated with battle operations. While the models generally estimate system effectiveness, based on various hardware and manning configurations, the scope and level of detail of these models vary tremendously.

For example, studies by McCallum and Underwood (1991) and Underwood, McCallum, O'Herin, Penner and Tyler (1991) illustrate how a spreadsheet can combine data about system information flows and task times with different combinations of automation. The data incorporated in this model are derived from subjective estimates of the frequency of event occurrence. By incorporating detailed task information (receiving information from databases and communicating specific intelligence information to other members of the staff), their model provides predictions of how the influence of automation on individual tasks contributes to the overall effectiveness of a military command and control center. The model output shows how system effectiveness changes as a function of the effect of automation on individual task times. More specifically, the output shows how automation affects the percentage of time operators are occupied. Thus, this model provides valuable data concerning the effectiveness of introducing automation. Failing to consider the dynamic nature of the task imposes one of the prime limitations on this modeling effort. This model does not reflect possible bottlenecks in the process caused by the dynamic nature of the system or the psychological limitations of the personnel. Another possible problem associated with the model lies in collecting data to enter into the model. The model relies on subjective estimates of changes in task time and therefore may not always reflect the actual task time. However, future versions may resolve these problems by using a more sophisticated modeling tool and validating subjective estimates with actual measurements.

A model developed by Corker, Cramer and Henry (1990) addresses some issues not encompassed by McCallum and Underwood (1991). This model of a command and control system also estimates the effect of automation on system performance; however, it incorporates a model of the cognitive and perceptual characteristics of personnel. In addition to estimating system performance, this model also estimates subjective workload, illustrating how the mental load varies across time. Thus, the output of this model allows designers to anticipate the effects of automation on the performance of command and control centers. Both of these models have been used by the military to anticipate the effects of various design decisions on system performance.

2.4 MILITARY WEAPONS SYSTEMS

Along with command and control systems, modeling techniques have been applied to a variety of military weapons systems. In particular, modeling efforts have investigated the effectiveness of weapon systems. For example, IDEAL (Chubb and Hoyland, 1988) represents a technique used to investigate the performance of a surface-to-air missile battery. The IDEAL methodology combines SADT (Structured Analysis and Design Technique) with SAINT (Systems Analysis of Integrated Networks of Tasks) to facilitate development of models of complex systems. SADT acts as a front-end analysis tool for SAINT, providing a structured approach to a top-down, hierarchical decomposition of systems that tracks the relationships between components. Since the description provided by SADT reveals only the static relationships between system components, a complete description requires a representation of performance over time. The SAINT portion of IDEAL incorporates the SADT description of the system into a model that represents the behavior of the system over time. Comparing the output of the model of a surface-to-air missile battery to empirical results, Chubb and Hoyland (1988) showed that the model succeeded in predicting the effects of implementing automatic tracking and variations in the types of sensors used to support manual tracking. Thus, the output of this model was used to guide system development. This model included the capabilities to provide a wide range of output data. In this application a key parameter was the time taken to respond. Other important parameters included elevation and azimuth tracking errors, and the probability of a hit. One of the primary reasons for the success of this modeling effort lay in the developers' ability to base parameters of the SAINT model on information drawn from a large body of empirical data. These empirical data included a detailed task analysis, estimates of task durations, and probabilities of errors. A similar maritime manning model might not have the same volume of data to draw upon. Other models, such as a model of howitzer crew drills (Bolin, Nicholson, and Smootz, 1990), demonstrate the continued success of these techniques.

While the studies of McCallum and Underwood (1991), Underwood, et al. (1991), Corker, Cramer and Henry (1990), and Chubb and Hoyland (1988) provide a detailed representation of a relatively small number of people, other models provide a much broader description of how the behavior and capabilities of many people may interact with a proposed system. For example, the Army uses HARDMAN III to deliver information on a wide variety of aspects of systems development. Composed of six modules it uses task network modeling (MicroSAINT) and data base access to generate data to guide system design. This system provides information about issues such as personnel availability, training requirements (given the current pool of available personnel), and effects of individual aptitudes on system performance. The scope of many of the analysis tools available through HARDMAN III exceeds the scope of interest for ship manning evaluations (Booher, 1990).

A Department of Defense initiative (MANPRINT) stipulates including human performance limits into military system design. To meet this requirement many tools have been suggested. In addition to the techniques already mentioned, a review by Booher (1990) describes a wide variety of models used to evaluate the effectiveness of military systems. These techniques range from models of a single person to models of large groups of people. The output of these models provides system designers with information upon which to base interface decisions, evaluate automation effectiveness, and develop appropriate manning structures. While these models can not be directly applied to resolve maritime manning decisions, they illustrate the feasibility of such an approach. Most of the modeling approaches used within MANPRINT are relatively simple techniques for assessing system user skills and requirements. As such, they do not permit predictions of system operation so much as serving as design-aiding tools.

2.5 PROCESS CONTROL INDUSTRIES

The process control industry has used simulation to evaluate the effect of technological innovations on manning and workplace design. Beville Engineering (1992) has developed a technique that allows them to compare staff workload, even though the contents of their jobs differ significantly. The data required for this analysis are based on job sampling, where intermittent observations of the workers provide snapshots of the workers' behavior and workload. Based on these data, "busyness" graphs provide equivalency curves representing the same level of "busyness" for different levels of time on tasks and tasks to be performed. Buck and Maltas (1979) used a more formal simulation technique, based on a task network model, to evaluate the difference in manning requirements of a hot strip steel rolling mill equipped with automation and another design that did not include sophisticated automation. The data for this type of model are derived from a detailed task analysis, based on observations and functional analysis. The output of this technique consists of timelines that illustrate operators' workloads and charts of the fluctuation of system performance over time. Examining this output they found that the simulation reflected the effects of technology found in the actual plants. These effects were based on a detailed validation process that compared the task times of selected mill operations. In addition, the model proved to be a useful decision making aid, enabling the users to estimate changes in overall system performance with various design alternatives, to calculate the time demands of various operator jobs, and to recognize interferences between tasks that might be alleviated with redesign. The authors concluded that modeling could be an effective tool for the evaluation of automation in process control plants.

Similarly, Munro, Martin, and Roberts (1968) used simulation to determine the optimum manning levels for continuous process control plants. This model succeeded in illustrating how manning structures should be adjusted as automation is introduced. More specifically, they successfully used the model to show how operators' workload would change with increased automation and different manning combinations. Although the model revealed differences between the highly automated process and the more traditional manufacturing methods, the authors noted several shortcomings of the model. Most importantly, they cited a difficulty in representing the cognitive loads and time demands associated with the tasks. Another shortcoming of the model was the difficulty associated with simulating a job that has never been performed. Both of these concerns apply to models that might be developed for the maritime industry. While this modeling effort identifies potential limits associated with developing models to estimate safe ship manning, it also illustrates the potential for success of such a model.

2.6 CONCLUSIONS

The history of models used to investigate the role of humans in domains other than the maritime industry illustrates the utility of models in assessing the safety and efficiency of complex systems. While the specific purposes and domain characteristics differ from the manning issues of concern to the Coast Guard, the generality of the approaches and the similarity of some characteristics of these domains make some of these approaches applicable to the maritime industry. More generally, the success of many of these models illustrates the feasibility of developing models to establish a technical basis for manning decisions. In addition, the limitations of these models illustrate some of the potential problems that a manning model may face. In general, the limitations of these models fall into two categories; (1) limitations of the modeling tools and, (2) limitations of the content incorporated into the model. In some instances the modeling tool or programming language used to create the model limits the interactions and characteristics of systems that the model can incorporate. For example, the spreadsheet style modeling tool used by McCallum and Underwood (1991) makes it impossible to incorporate dynamic characteristics of the system, such as transient changes in task frequency that lead to performance fluctuations. In addition, content or data included in a model may limit its utility. For example, CES provides an accurate representation of operators' diagnostic ability, but only for a few types of scenarios. This examination of the tools and content of previous modeling efforts illustrates the feasibility of developing models to predict performance of human-machine systems. It also reveals some of the limits of modeling tools and content that should be considered in any modeling effort undertaken by the Coast Guard.

III. FEATURES OF TOOLS AND CONTENT USED IN MANNING MODELS

Chapter II showed that modeling often allows other industries to predict the performance of human-machine systems, providing a technical basis for decisions involving interface design, automation implementation, and manning requirements. While the usefulness of models in resolving specific issues for which they have been designed has been demonstrated, none of these models could be readily adapted to provide a technical basis for all manning decisions in the maritime industry. These models are limited by the tools used to construct them and the information they incorporate. Just as each of the models developed for other domains is limited by the tools and information used in their construction, models developed to evaluate safe manning practices will also depend on these two factors. If either the characteristics of the modeling tool or the model content is ignored, then the model may not provide a valid technical basis for manning decisions.

3.1 FEATURES OF MODELING TOOLS

The brief review of modeling approaches used in other domains illustrates the great diversity of potential models. Many different tools could represent a ship system, each providing a different method of translating the characteristics of the ship system into a model that can provide information about the consequences of various manning approaches (e.g., increased automation, reduced manning, and emergency response capabilities). The variety of modeling techniques ranges from simple spreadsheet calculations to complex approaches that combine the state-space representation of modern control theory with artificial intelligence pattern recognition. Each one of these tools offers different capabilities for modeling the manning requirements of a ship. Likewise, each tool requires different resource requirements (e.g., development time and expertise) for model development and use. Therefore, the development of a valid and economic model depends on balancing model capabilities with its resource requirements.

Differences in tool characteristics can have a substantial effect model validity. For example, one difference between the tool used by Argent Marine and that used by the NRC is that the Argent Marine model predicts peak periods of overload that the NRC model would ignore. Thus, the characteristics of the modeling tool used to create the manning model can lead to underestimating ship personnel requirements. Because limitations of modeling tools can generate inappropriate recommendations for ship manning, it is crucially important to match the capabilities and limitations of the modeling tool to the characteristics of the system to be modeled. Modeling tools contribute to the validity of a manning model by limiting the types of and nature of the relationships that a model can include. If the model capabilities fail to match the characteristics of the system being modeled then model validity will suffer.

Some prototypical modeling tools include: spreadsheet, flow diagram, event tree, timeline, task network, control theoretic, queuing theory, and artificial intelligence. Table 1 summarizes the limits and capabilities of a variety of modeling tools; these are discussed in more detail in Appendix A. Because of the diversity of features available in modeling tools, no one tool contains all possible features. Some tools contain many different features and others contain relatively few. The techniques that include more filled dots in Table 1 offer the capability of modeling a wider variety of systems; however, these differences do not necessarily imply the superiority of any one tool. For example, the task network approach supports a wide variety of modeling capabilities, but the diversity of capabilities alone does not guarantee success. Specifically, the timeline approach has far fewer capabilities, but it includes the ability to model continuous variables, which could be essential in modeling some systems. This comparison shows that the number of dots on Table 1 alone is not a good metric of a modeling tool. Instead, the worth of a tool depends on how closely tool characteristics match system characteristics. Therefore, the choice of a tool depends on the match between the capabilities of the tool, the characteristics of the system, and the types of decisions a model is meant to support.

For example, to evaluate whether a proposed crew configuration will satisfy workhour regulations does not require an extremely sophisticated modeling tool. Instead of developing a complicated task network representation of the ship system, a more simple representation, based on a combination of the timeline and spreadsheet tools, would suffice. Although a task network representation captures the probabilistic and contingent nature of shipboard activity, these subtle details may not have a significant influence on the ability of the proposed crew to meet workhour limitations. Likewise, an artificial intelligence-based model might accurately portray the decision making of ship personnel, but this level of sophistication may have little effect on the estimation of workhour requirements.

To support matching system characteristics and tool capabilities, Table 1, and the more detailed discussion included in Appendix A, can be used as a checklist to verify that the tool can model the system. Appendix A provides a classification of modeling tools to illustrate their various capabilities, followed by a comparison of the capabilities of these tools based on the classification. The comparison of different modeling tools illustrates how their capabilities and limitations influence the ability of a model to represent the ship system accurately.

In addition to the implications for model validity, the characteristics of modeling tools also influence the resources needed for model construction. Specifically, differences in tool complexity lead to differences in the expertise and time required to develop and operate a model. Some tools offer simple intuitive interfaces that enable novices to enter data and create a model. Other tools require much specialized knowledge. The spreadsheet, flow diagrams, event trees and timelines require little training for model development. These tools require only a thorough knowledge of the specific system, a few modeling formalisms, and a knowledge of the limits of the tools. Task network modeling approaches require substantial knowledge of the specific tool, comparable to learning a programming language. Models relying on control theoretic, queuing theory, and artificial intelligence all require substantial expertise related to the theories underlying

Table 1. The relationship between modeling tool capabilities of modeling tools and tool characteristics.

Tool Characteristics / Tool Types	Spread-sheet	Flow Diagrams	Event Trees	Timeline	Task Network	Control Theoretic	Queuing Theory	Artificial Intelli-gence
VARIABLE REPRESENTATION								
Continuous				●		●		○
Discrete	●	●	●	●	●		●	●
Combined				●				○
TASK ORDERING								
Concurrency		●		●	●		●	
Priority					●		●	●
Deterministic	●	●	●	●	●		●	●
Contingent		●	●		●	●	●	●
Probabilistic			●		●	●	●	○
EVENT TIMING								
Deterministic	●		●	●	●		●	○
Contingent			●		●	●		○
Probabilistic			○		●	●	●	○
RESOURCE REQUIREMENTS								
Deterministic	●	●		●	●	○	●	
Contingent					●	○		
Probabilistic					●	○	●	
RELATION TO SYSTEM								
Task sequence	●	●	●	●	●		●	●
Open loop			○		●	●	●	
Closed loop					○	●		

● Tool accommodates indicated dimension.

○ Tool marginally accommodates indicated dimension.

these approaches. In addition, the complexity of these approaches requires much more development time, compared to other approaches. Thus, the order of modeling tools in Table 1 provides a rough metric of the expertise and time required to create a model with each of these approaches. In other words, by presenting the models in a left to right sequence, starting with the spreadsheet and ending with the artificial intelligence approach, Table 1 indicates expertise and time required to use each of these tools. Although tools at the right of this table can be used in a wide variety of circumstances, their extra capabilities require substantially more resources than the more simple modeling tools. Conversely, while tools with many features may be cumbersome for investigating a simple system, a simple tool may not contain the features required to represent a complex system accurately.

3.2 FEATURES OF MODEL CONTENT

While the classification of modeling tools illustrates how limitations of tools contribute to the usefulness of manning models, the information incorporated into a model, regardless of technique, also determines the type of data that the model requires and the type of information the model will produce. Similar to the great variety of modeling techniques, a model of a ship system may incorporate many different types of information. The content of a model may range from a simple description, limited to a small part of a ship, to a very detailed description that encompasses an entire ship. As with the different modeling techniques, different types of information incorporated into the model will influence both the validity and cost of the model. Therefore, model validity and resource requirements depend not only on the modeling tool, but also on the model content.

In general, a model becomes more accurate as the content of the model becomes more inclusive, matching the actual system more closely. However, the more information the model contains, the more complex and costly it becomes. Since no model perfectly replicates the actual system, the content of a manning model should depend on the specific questions that the model should address. Appendix B provides a classification, based on six dimensions, that describes the different types of information that a model might incorporate. Each of these six dimensions (scope, level of detail, interactions among subsystems, level of abstraction, veridicality of content, and organization of model content) can describe each of the two primary components of a manning model (see Appendix B). The first of these components is a description of **task demands** (what personnel must do) imposed on ship personnel. These task demands depend on the particular trade route, ship configuration, maintenance program, level of automation, or emergency situation. The second component of a manning model is a **profile of factors** that relates the task demands to the safety, efficiency, and legal constraints associated with operating a ship. These factors might include human performance limitations such as fatigue, memory and strength, as well as the effects of high mental workload and the limits of decision making effectiveness. In addition to human performance limits, a manning model might also include factors that relate task demands to the rules, regulations, and guidelines that govern the maritime community. For example, these factors might show that task demands require the crew to violate the 12 hour work day rule imposed by the Oil Pollution Act of 1990. A single model might include both human performance and regulatory limitations; however, constructing two models, one based on human performance and the other based solely on the legal constraints, could provide a clear illustration of regulatory deficiencies. While all six content dimensions apply to description of both the task demands and the profile of factors, some dimensions are more important in the description of task demands, and others are more important for the description of the profile of factors. In summary, each content dimension can describe the task demands imposed on ship personnel, as well as the profile of factors that overlay these demands to reveal the legal, safety, and efficiency consequences for the system.

Just as characteristics of tools contribute to the capabilities of manning models, model content contributes to the ability of a model to mimic reality. The detailed description of six dimensions of model content included in Appendix B shows how each of these dimensions affects model accuracy. Some of these dimensions (scope of analysis and level of detail) refer to the breadth of a model's description of a system. Other dimensions, such as the interactions among subsystems, level of abstraction, and veridicality of content, refer to the types of relationships included in a model. The final dimension, organization of model content, illustrates the limits of the model based on how information is incorporated into the model. This organization specifies which variables may be manipulated and which are fixed. These six dimensions of model content combine with the limits and capabilities of a modeling tool to determine the validity and type of information that a model can produce. Therefore, creating a manning model requires special attention to the types of information included (or available) in a model and its organization. Appendix B provides a guide for this process, showing how information and its organization influences model validity and costs.

3.3 CONCLUSIONS

This brief review shows that characteristics of modeling tools and model content play a major role in the accuracy of the manning requirements generated by a manning model. If either the tool or content fails to match the characteristics of the system to be modeled, then the manning requirements produced by the model may seriously underestimate the actual manning requirements. To avoid dangerous situations produced by inadequate manning, the validity of the model needs to be evaluated. The classifications of modeling tools and model content support an analysis of the match between the modeling tool and model content and the characteristics of the actual system.

IV. COMPARISON OF EXISTING APPROACHES TO MODELING MARITIME MANNING

Of the many models developed for the maritime domain, some address the issue of manning, while others concentrate on more specific issues, such as the effect of new navigation systems on crew performance. Other models focus on hardware requirements and reliability not directly related to ship manning (e.g., a port simulation model of bulk cargo operations by Park and Noh, 1987). Since a relatively large number of models exist, this chapter contains a brief summary of several models, followed by a more detailed analysis of six models that have particular relevance to developing a manning tool.

One of the modeling efforts closely related to ship manning is a report by the PRC (1992) that describes some of the issues involved with manning decisions, and provides an outline of a manning model. Their report distinguishes between "manpower requirements" (limits on manning levels imposed by task demands) and "manning level" (imposed by regulatory requirements). Their objective is to determine "manpower requirements" with different levels of automation. To achieve this goal they have developed a manpower/technology matrix that illustrates how increased levels of automation might translate into lower manning levels; however, the current PRC report provides little substantiation for these claims. Their proposed model consists of a spreadsheet representation of a functional task analysis that shares many similarities with the task analysis included in the NRC model. The spreadsheet model simply sums task times assigned to each crew member; excessive time demands indicate inadequate manning. The estimated task times and personnel assignments incorporated in the model will be gathered through an industry questionnaire. Overall, the current report provides little indication of how this model would support a technical estimate of the manpower requirements as a function of different levels of automation.

In addition to its imprecise description of the effect of automation on manning requirements, the PRC report fails to incorporate human performance limitations and their effect on ship safety. Perse, Baker, and Malone (1990); Perse, Callahan, and Malone (Unpublished contractor's report), and Malone, Baker, and Permenter (Unpublished contractor's report) illustrated the development of models that address human performance limitations and their effect on manning requirements. For these applications, the models provided estimates of manning requirements and design improvements based on workload and information processing bottlenecks. Specific applications addressed emergency response to flooding and fire on a ship. Specifically, the Navy has used one of their models (Perse, Callahan, and Malone, Unpublished contractor's report) to identify information processing bottlenecks, operator workloads, and performance capabilities of several scenarios associated with an Integrated Survivability Management System (ISMS) damage control system on an Aegis-class ship. By examining the effect of a variety of parameters (automation capabilities, personnel allocation, task priorities, and task time) the model provided a basis for evaluating the changes to damage control doctrine, procedures, personnel allocations and design options (Perse, Callahan, and Malone, Unpublished contractor's report). These models have been developed with SIMWAM (Simulation for Workload Assessment and Modeling), a modeling tool that provides the full functionality of the task network modeling technique. These features include: probabilistic, deterministic, or conditional task ordering, with interruptions based on priority. The successful use of this tool in specific maritime applications illustrates the utility of the task network approach to modeling, as well as the value of incorporating human performance limits and workload estimates.

Perse, Baker, and Malone (1990) incorporate human performance limits based on characteristics of individual tasks; however, many of the tasks that occupy mariners are not discrete, but continuous. For example, manning the helm requires continuous adjustments when executing a series of maneuvers under manual control. To examine this type of situation Veldhuyzen and Stassen (1977) adopted a control theoretic model to describe the influence of additional information (e.g., rudder angle and rate of change of rudder angle) on the performance of the helmsman. This model accurately predicts both the helmsmans' control actions and the deviation of the ship from the desired course. While this model provides an excellent example of incorporating humans' perceptual and motor limitations to predict ship performance, it is not a manning model, and provides information more suited to design decisions, with little relevance to determining safe manning levels. However, design changes may drive demands for manning reductions, and a model that illustrates the specific effects of design changes on the ability of the crew to control the ship may provide a technical basis for interpreting the effects of technology on manning requirements.

In addition to the models developed by the PRC, Malone et al., and Veldhuyzen and Stassen several other models exist, which deserve closer inspection. These models include: (1) Analysis of a collision-avoidance system with operational sequence diagrams, (2) Design for maintainability with modified petri nets, (3) Integrated system reliability - a model of overall human-machine reliability, (4) Maritime Administration manning model - an analysis of selected ship functions based on multiple activity charts, (5) NRC functional manning model - an inclusive spreadsheet model of manning requirements), and (6) Argent Marine manning model - a relatively sophisticated manning model based on a project management tool. These six models represent a variety of modeling techniques, each containing quite different information. With the exception of the NRC model, these six models were not developed as generic manning models. Instead, they were developed for specific purposes. Thus, the models may have limits with respect to serving as a comprehensive manning model even though they may have successfully fulfilled the purpose for which they were designed. An examination of these models illustrates their ability to meet the objectives for which they were designed. In addition it illustrates the limits of each approach, relative to the development of a technical basis for ship manning. More specifically, the discussion addresses several aspects of the models including: the purpose, the type of tool, the content, the output, the analysis of its capabilities, and the recommendations for future application.

4.1 ANALYSIS OF A COLLISION-AVOIDANCE SYSTEM WITH OPERATIONAL SEQUENCE DIAGRAMS

Purpose of model: This model evaluated the ability of a shipboard radar computer to improve navigation safety by reducing the likelihood of ship collisions. An examination of the elements of the collision avoidance task provides the basis for a comparison of the old and new systems, revealing the effect of new technology. Kurke (1961) illustrated how an Operational Sequence Diagram (OSD) model could provide information about timing of events, probabilities of success, and location of activities.

Modeling tool: When Kurke (1961) introduced the OSD technique, he illustrated the technique with a model of two collision avoidance systems. The OSD technique g phically illustrated the information gathering, decision, and action sequences involved with the collision avoidance task. The OSD method belongs to the general class of flow diagram techniques. Table 1 in Chapter III, illustrates many of the capabilities and limitations of this class of techniques. As a flow analysis technique, OSD primarily represents the system as a deterministic flow of events (a simple task sequence). Each major decision point generates alternate action sequences, and requires a new diagram. Therefore, situations involving many decision points will lead to complicated sets of diagrams. Related to these limits, the OSD approach does not provide for contingent or probabilistic event timing, i.e., all events or tasks have a predefined duration. The graphical representation of the OSD can take three forms. First, it can represent the time dependent interaction between the radar, ship, watch officer, and computer. Second, the OSD can convey the formal logic of the task sequence, showing the precedence and ordering requirements of the tasks. Third, the OSD can represent the spatial relationships between the radar, ship, watch officer, and computer. Thus, the OSD technique can represent interactions of system components and information flow across time, the logic implied by the task constraints, or the spatial distribution of activity.

Model content: OSDs require a complete description of the task sequence of the system being examined. Each different OSD representation (interactions across time, logical constraints, and spatial distribution of activity) requires different types of information. Kurke (1961) developed three different OSD models to illustrate how these different representations could be used to evaluate human-machine systems. A comparison of the interactions across time required estimates of time involved with the actions and decisions. The analysis of the logical constraints required the probabilities of decisions and actions, together with their consequences. An analysis of the spatial distribution of activity required overlaying activities represented in the OSD upon a diagram representing the physical location of the various panels that comprised the human-machine interface. Thus, the model presented by Kurke (1961) included a description of the task sequences, together with task times, event probabilities, and spatial distribution of activity. This information was gathered from observations of personnel and analyses of procedures.

Model output: The output of the model described by Kurke (1961) took three forms. First, the graphical representation of the tasks across time illustrated the decisions required, and provided a qualitative estimate (only relative duration, with no exact numerical representation) of the duration of the overall task and the activity level of personnel across time. Second, the representation of the logical constraints of the task provided an estimate of the probability of various task outcomes. Third, the representation of the spatial distribution of tasks revealed how the physical layout of the interface supported the specific task under examination. These three types of output were used to identify possible design changes to prevent errors from occurring, function allocation tradeoffs based on the relative reliability of the automation and humans, and interface design based on predicted patterns of information use. Thus, the models developed by Kurke showed that a variety of models might be needed to address issues involving safe ship manning.

Analysis: While the OSD technique illustrated by Kurke (1961) provides a graphical representation of the task sequence and the interaction with system components over time and space, as well as a framework for estimating the probabilities of events, the technique has several limitations. For example, OSD provides a qualitative representation of event timing and task loading of personnel; however, it is not a convenient method to assess quantitatively either event timing or task loading because this information must be inferred from the diagrams and manually summarized. Regarding human error probabilities, the OSD technique fails to accommodate information concerning factors that might influence the probability of errors. Specifically, changes in task loading or time stress do not influence the outcome of decisions or actions, whereas they greatly influence decisions and actions in actual situations. One of the primary limits of this technique is the complexity produced by decision branches. With the paper-based approach to OSD, a new diagram begins at every major decision point. However, a computer-based equivalent might facilitate integrating these multiple diagrams more readily than a collection of paper-based diagrams. Even so, this technique does not include formalisms that support contingent and probabilistic branching.

While the OSD technique and models described by Kurke have several limitations, they also offer potential. The output of the models developed by Kurke demonstrate the feasibility of developing a formal representation of human behavior. Specifically, his models show how OSDs might be useful in understanding decision making, the effect of human error, and the spatial distribution of tasks. Based on this demonstration, OSDs could be developed to address issues such as the effects of reduced manning on ship reliability and the ability of bridge design to enable manning reductions.

Recommendations: The OSD technique provides qualitative information about situations where a sequence of information transmission, control actions, and decision making events define system activity. The limitations of this technique in estimating manning requirements make it most useful in estimating personnel requirements for specific, short term tasks, such as docking, which has a relatively well defined and predictable task sequence.

4.2 DESIGN FOR MAINTAINABILITY WITH MODIFIED PETRI NETS

Purpose of model: The study conducted by Madni, Chu, Purcell, and Brenner (1984) developed a generic approach to modeling and analyzing shipboard multi-person maintenance tasks with the aim of identifying and alleviating potential human related maintainability problems. This model also provided estimates of workload levels, which have potential for verifying the adequacy of manning levels.

Modeling tool: To address maintainability of equipment, Madni, Chu, Purcell, and Brenner (1984) adopted a "Modified Petri Net" modeling tool. Belonging to the category of task network modeling techniques, it incorporated only discrete events, with no representation of continuous variables. However, this technique was particularly useful in modeling both asynchronous and concurrent activities, as well as constraints on their precedence, initiation, resource requirements, and frequency. Using a discrete event representation, this tool facilitated an analysis of how individual tasks, events, and contingencies (related to the environment or to human activity) affect system performance. As a network-based modeling approach, it provided a representation of the dynamic flow of information, which illustrated the temporal ordering of tasks and the propagation of the effects of those tasks over time. Also, this modeling tool facilitated construction of a hierarchical model, where tasks included subtasks and sub-subtasks as the analysis demanded a greater level of detail. The modeling tool facilitated a representation of the system based on abstract concepts of information flow, as well as more concrete issues such as task duration.

Model content: The scope of this model included a description of the decision-making sequence involved with the lubrication of the main reduction gear, and the activities of the engineering officer on watch and the other engine room personnel. The simulated activities depended on a normative model of the maintenance tasks, constructed by examining existing procedures and collecting opinions of domain experts. The model incorporated human performance limitations by including workload estimates based on whether the activity demanded skill, rule, or knowledge-based performance (Rasmussen, 1986). Those tasks that require only skill-based performance imposed the least workload demands, while those requiring knowledge-based performance imposed the highest workload demands.

Model output: The output of the model included mental workload measures, time estimates, and errors affecting procedure execution. The provided estimates of instantaneous and cumulative mental workload, where the instantaneous workload related to the stress imposed by immediate tasks, and the cumulative workload related to fatigue resulting from a series of tasks. The model represented instantaneous workload as the summation of task-related workload and event-related workload for tasks and events currently active, and it represented cumulative workload by integrating instantaneous workload over time. The model also produced estimates of the workload associated with a variety of paths through the task network (specific series of tasks). This feature provided an estimate of the worst case workload situation, where a human must perform a number of highly demanding tasks. Thus, the output of this model provided an estimate of whether human capabilities match the demands of maintenance task, where mismatches signal the need to modify the maintenance task or increase the capabilities of the crew through training, aids, or additional manpower. These results illustrate that modeling can provide a powerful method of estimating the ability of ship personnel to maintain a ship effectively.

Analysis: The model was developed to examine maintenance activities, and in this role it was a success. It showed how workload demands of maintenance tasks vary for different decision making sequences. From the perspective of using the model to evaluate ship manning, the primary limitations of the model stem from its limited scope. Since the model only examined maintenance activities associated with the lubrication of the main reduction gear, the model provides little insight into overall vessel manning. Additionally, the model is not specifically designed to evaluate manning (the measures and framework were not designed to compare different manning structures). The model output supported an analysis of changes to procedures and hardware configuration. However, the sophistication of the modeling tool employed facilitates a complex and realistic portrayal of maintenance tasks that, once enhanced, could help resolve manning issues related to maintenance requirements.

Recommendations: This model illustrates an application of a sophisticated modeling technique that incorporates task demands and human performance limits to produce workload estimates and other task related performance measures. The development of a model to evaluate ship manning should incorporate some of the techniques illustrated in this model, such as workload estimation, the effect of task concurrency, and the type of mental functions required for tasks. More specifically, this model may be best suited for application as a specific sub-model, such as in the evaluation of maintenance capabilities.

4.3 INTEGRATED SYSTEM RELIABILITY

Purpose of model: Siegel and Wolf developed a large number of man-machine models that have been applied to single person and large group situations, with crews up to 90 men (Siegel and Wolf, 1969). Siegel and Wolf (1974) presented a model of intermediate size crews (4-20 men). The primary aim of this and other models produced by these authors was to integrate human reliability with equipment reliability to estimate the effect of human performance limits on the effectiveness of naval vessels. As such, these models do not specifically address manning considerations, but include manning with a wide variety of variables that affect ship reliability.

Modeling tool: The model was developed using FORTRAN IV language to implement flow charts that resembled a task network modeling approach. As such, the tool used to develop this model included both the benefits and limitations of the task network techniques shown in Table 1. More specifically, this framework supported the development of models that could include missions with durations of up to 30 days. Each mission could include 300 event types, 200 of which could be scheduled on any given day. Each of these events could involve up to 20 men who represented 10 specialty areas at any one of 10 command levels. The modeling framework also provided the ability to develop a model that included up to 30 types of equipment and 10 types of emergencies.

Model content: The content of the model fell into three areas: properties of human performance, characteristics of the mission, and attributes of the equipment. Human performance factors included a wide range of variables, such as the effects of motion sickness, fatigue, physical incapacitation, competence, mental and physical workload, and goal aspiration. Previous research provided the basis for estimating the effect of these variables on human performance. Mission variables included sea state, fixed and variable task times, task repetition if poorly performed, and task precedence. Equipment-related variables included: equipment reliability, the opportunity for human initiated failures, and performance levels.

Model output: The complexity of the model provided a wide variety of output measures that Siegel and Wolf placed into broad groups consisting of mission effectiveness, time utilization, and personnel. Mission effectiveness included the more specific measures of system reliability, system performance, estimates of equipment and human reliability using mean time between failures (MTBF) and mean time to repair (MTTR). The category of time utilization included variables such as: task success rates, idle time, sleep, repair, and the number of tasks ignored. The personnel category included variables such as: physical and mental load, performance adequacy, and health and safety indices. The model generated these measures for each simulated event, day, mission, or in a summary of the simulation run that included as many as 30 days. The output of this model provides a quantification of the reliability of man machine systems. More generally, it provides a vehicle for system planning in the context of a variety of personnel constraints. For example, this model illustrates the effect on system reliability of different crew sizes and workday lengths.

Analysis: Although this model has been successfully used by the Navy to evaluate ship reliability, it has several limitations if it was to be applied as a manning model. The primary limits in applying this model to the manning issues of the Coast Guard lie in the ease of modification, the need for further validation, and the need to define events. The implementation of the model described by Siegel and Wolf (1974) failed to provide an interface that facilitates easy access to either the model parameters or the event structure. Therefore, modifications to the simulation could be cumbersome. In addition, many of the aspects of the model require further validation and elaboration to represent the manning requirements of a ship accurately. For instance, the relationships that govern fatigue, depend solely on the time period immediately preceding the work period (8 hours of sleep leads to a fully recovered seaman) and does not consider the cumulative effects of sleep disruption. Another requirement of this model is the need to complete a detailed task analysis to define the events and the parameters associated with the events. The NRC manning model provided a first step in this process, and the Siegel and Wolf modeling approach could provide a framework for incorporating information gathered in a more detailed task analysis. The success of this and similar models used by the Navy to investigate the effect of human behavior on system performance shows the viability of this technique, and of modeling in general.

Recommendations: With further validation and modification to suit the specific needs of the maritime industry, many of the issues addressed in the Siegel and Wolf model could form the basis of a manning model. The Siegel and Wolf framework is especially valuable because it provides measures of not only time demands of shipboard tasks, but it also provides measures of how those time demands might interact with human performance variables such as fatigue, workload, and stress to produce an estimate of overall system reliability. Constructing a manning model that includes the capabilities demonstrated by Siegel and Wolf (1974) would require extensive data collection using surveys, direct observations, and expert judgment. The extensive data collection that would be needed to address the wide variety of factors affecting manning decisions would make the development and use of a manning model similar to that of Siegel and Wolf (1974) extremely expensive.

4.4 MARITIME ADMINISTRATION MANNING MODEL

Purpose of model: The purpose of this study was to investigate the manning requirements and skills necessary to operate and maintain vessels built after 1960 and before 1970 (U.S. Maritime Administration, 1983; Williams, 1983). Three types of vessels were studied; they included a container ship, a LASH, and a Sea Bee (a barge carrier). The engineering plant for each of these vessels consisted of a steam turbine plant with a self-regulated boiler, automatic boiler alarms, and safety shutdown devices.

Modeling tool: The modeling tool adopted for this manning model consisted of a paper-based "Multiple Activity Chart", resembling a timeline technique, with limitations and capabilities illustrated in Table 1 of Chapter III. Specifically, the Multiple Activity Chart technique provides the ability to illustrate the occurrence of concurrent events; however, this technique does not provide quantitative estimates of time and resource requirements. It does provide a deterministic sequence of events. The particular Multiple Activity Charts developed for this model included discrete tasks and contained no continuous variables such as fatigue, workload, or system performance. The model consisted of charts with time running vertically down the page, and with columns for each person involved with general functions (bridge watch or anchoring). The column associated with each crew member contained blocks of time devoted to specific tasks ("act as lookout" or "maintain course"), and were aligned with the timeline to illustrate when specific tasks began and finished.

Model content: The manning model presented in the report contains task information at a fairly fine level of detail. For example, the description of the "Bridge watch" specifies tasks such as "Prepare weather report", "Check RPM", and "Scan horizon". While functions such as "Bridge watch" contain much detail in the listing of specific tasks, the exact time requirements and the sequential dependencies remain unclear. In addition, the model provides little information concerning the more global aspects of ships manning. For example, the Multiple Activity Chart constructed for bridge watches covers only the 1600-2000 and 2000-2400 watches, providing little indication of the structure of crew rotations and the overall number of hours of work allocated to each person.

Model output: The model output was a qualitative description of tasks and their ordering, performed by each crew member to satisfy each shipboard function. The description of the model provides no quantitative measures of workload, time requirements of tasks, or system performance (Maritime Administration, 1983). The task listing presented in the Multiple Activity Chart provided a qualitative estimate of the relative workload levels and the spare capacity of crew members. The models developed in the Maritime Administration study illustrate how even a very simple model might aid in developing a technical basis for ship manning by organizing the data into a more comprehensible form.

Analysis: The primary limitations of the model stem from the type of modeling tool and the limited scope of the information included in the model. The Multiple Activity Chart technique provided only qualitative information concerning task ordering and relative workloads; however, these could be deceiving if task times were inaccurately plotted or task assignments failed to reflect the true organization of the crew. For example, the master and mate might share responsibility for several tasks, and when one becomes overloaded the other accepts more tasks. Because a timeline approach fails to represent contingent task ordering, this model did not represent this sort of workload management accurately. In general, this modeling technique was subject to the limits of the timeline technique (See Appendix A). The limited scope of this model did not address issues related to shift coordination and emergencies. Models produced by Argent Marine and the NRC represent entire voyages (see below), whereas this model represents only major ship functions such as "Bridge Watch", "Catering", and "Engineering Plant".

This model succeeded in achieving its goal of projecting a reasonable crew size and composition for the three types of vessels it examined. While the modeling technique used in this study has several limitations (with respect to serving as a generic manning model), it provides a method of examining the relationships between concurrent activities. This capability facilitates the identification of critical situations that might require coordination of several personnel. Many other techniques ignore such situations, which could lead to underestimates of manning requirements. Thus, this technique can provide valuable insight into manning requirements that other techniques would not.

Recommendations: While the specific manning model presented in this paper provides little quantitative evidence for estimating manning requirements, the qualitative analysis that accompanies this model provides an excellent basis for future modeling efforts. The Multiple Activity Chart representation supports only a qualitative representation of the system, limiting it to a supplemental role in future manning models.

4.5 NRC FUNCTIONAL MANNING MODEL

Purpose of model: The purpose of the model was to support an assessment of manning applicable to all ship types, classes, and trades. In addition, the model provided an assessment of manning requirements over all voyage phases, as well as emergencies.

Modeling tool: The description of the model included in the 1990 NRC report suggests that the committee employed a spreadsheet type modeling tool. This modeling tool combined the frequency of tasks identified in the functional analysis, with the crew and time requirements for each task to generate a total number of hours of crew time that a voyage requires. The spreadsheet type tool provided quantitative estimates of time requirements; however, it failed to capture the sequential nature of tasks, the effect of task ordering on the crew's workload, or the effects of concurrent tasks on performance. Table 1, in Chapter III, illustrates other deficiencies of the spreadsheet-based modeling approach.

Model content: The scope of the model encompassed multiple voyage phases, emergency situations, and ship types. While the model contained no information concerning human performance limitations it did contain estimates of the frequency and duration of a wide variety of tasks. The level of detail of these tasks was at the function level; it did not consider the tasks that comprised this function. For example, this model included tasks such as: "cleaning/wash down of deck", "vessel fabric maintenance", "helicopter operations", "inert gas operations", and "inert gas unscheduled maintenance". The time and expertise requirements for these tasks were obtained via questionnaires and expert judgment.

Model output: The spreadsheet tool limited the output of this model to a simple summation of the number of hours required by each of the different skill categories. For example, the model might have specified that the ship required 282 hours of work for personnel working in the steward's department. The output gave no indication of the distribution across time, or across individuals within a skill class. However, the model did provide a large catalog of tasks and their time requirements. This catalog, together with some simple spreadsheet calculations, provided a rough estimate of the manning requirements for a ship. While this model lacked many refinements that would lead to more accurate manning estimates, it showed that this type of approach could help support technical estimates of crew size.

Analysis: The limitations of the NRC manning model stem from the characteristics of both the modeling tool and the model content. The spreadsheet modeling tool prevented the model from revealing possible concurrent activities that would require the simultaneous presence of a single person at different parts of the ship. Likewise, the modeling tool prevented an analysis of the distribution of workload across personnel and time. For instance, the model could not differentiate between a task requiring two persons for two hours or one person for four hours. Similarly, the NRC model would have predicted a safe manning level with a task loading of 280 hours per person over a six week voyage; however, if 200 hours of the 280 occur during port operations (which occupy a total of two weeks of the six), then the scheduled workhours would exceed 14 hours a day.

The content of the NRC model also limited its usefulness. One of the primary limits of the model involved the level of detail of the tasks. In many cases, even the most detailed description of the task failed to specify the nature and duration of the task adequately. For example, unscheduled maintenance of the evaporators might require a wide range of times depending on the extent of the problem. Simply selecting an average task time of several hours may significantly misrepresent a task that occurs infrequently, but when it occurs requires many hours of attention. A similar problem exists for all the task time estimates included in the model. All task time estimates fail to include any information about the distribution of times, so the model represented tasks that always take five minutes to complete in the same way as tasks that sometimes take one minute, but at other times may take as long as 20 minutes (averaging five minutes). Not only did the model fail to provide the ability to estimate effects of task time variability, it provided no indication of how human performance limitations influence system performance and task times. Another problem with the data included in the model was the vague description of the methods used to collect it. Unless the method adopted to estimate task times provided accurate estimates, the output of the model would have been of little worth. Besides problems with data concerning task times and frequencies, the model contained no information (such as error rates or probability of machinery failures) required to support a reliability analysis.

While this model has several limitations, it has the ability to support a quantitative analysis of several manning issues. For example, once data for a given voyage profile is collected, the model can provide gross estimates of the effect of new technology or revised operating conditions on manning requirements. Additionally, this model provides information regarding the ability of crew to respond to emergencies and planned activities such as port operations. For example, in their validation study the NRC demonstrated the ability of the model to estimate the number of people required to man the ship during a variety of fires, including container fires and fires caused by high-pressure fuel line breaks. These estimates could aid in identifying manning-sensitive and manning-rich tasks (NRC, 1990). The documentation of the NRC model makes it difficult to determine how well it would support this type of analysis for ships other than the two types mentioned in the report. However, the documentation suggests that detailed data collection, based on expert opinion or observations, would be required.

Recommendations: The limitations associated with this model do not imply that it is not a useful tool for examining manning levels. As long as the assumptions and limitations are consistent with the nature of the system and the goals of the user, the model may be quite useful in judging manning levels. For example, in their validation study, the NRC was able to use the model to estimate the crew requirements (number and organization) of an American President Lines C-9 cargo ship. The model also showed that under different operating conditions these manning requirements could be reduced (NRC, 1990; pp. 70). The validation study presented in the NRC study suggests that the model (when used by its creators) represents the tasks associated with operating a ship closely enough to provide reasonable estimates of manning levels. The ability of others to generate similar data would further enhance the validity of this model; however, the sometimes vague description of terms in the task analysis might require expanded definitions.

4.6 ARGENT MARINE MANNING MODEL

Purpose of model: Argent Marine is currently developing a model that describes the tasks and the associated time demands for LNG tankers traveling between the Middle East and the east coast of the U.S. This model provides data for only these ships on this trade route.

Modeling tool: The modeling tool adopted by Argent Marine is a project management software tool. In many ways the project management software resembles a timeline approach; however, it has significant differences that substantially increase its capabilities. For example, it monitors the assignment of resources to tasks, and identifies situations where multiple tasks require the same resource. For example, the software would automatically indicate a lack of resources if the chief mate must simultaneously man the bridge and prepare for cargo transfer. In addition, the project management tool also facilitates modeling contingent events. For example, a model could include the fact that a ship must be docked before cargo operations may begin. Besides these differences, the project management approach represents the system as the timeline approach, containing the capabilities of this technique as shown in Table 1, Chapter III.

Model content: The content of the model includes a hierarchical task listing similar to the one included in the NRC model. Like the NRC model, many of the task descriptions do not contain very much detail with tasks such as, "Watch-standing", "Docking", and "Planned Maintenance on the Anchor Windlass" representing the types of task descriptions incorporated in the model. For every task, the model includes task time and crew requirements, based on expert judgment. The model does not explicitly include any human performance limitations, such as fatigue, mental or physical workload, or constraints on decision making.

Model output: The output of the model includes a variety of graphical and numeric reports. These reports consist of such features as timelines that illustrate the allocation of specific personnel over any specified time. The model can also provide illustrations of the total hours worked over any specified time. In addition, the software produces Gantt charts that show the sequential relationship between tasks, illustrating how the completion of some tasks depends on the completion of others. Even though this model has a variety of sophisticated features it does not provide any data that relate task demands to ship performance or reliability. While this model has several weaknesses, it represents a sophisticated method of calculating crew requirements. The capabilities of the project management software enable the model to estimate the dynamic changes in the demands on the crew. This provides more accurate estimates of crew requirements than are possible with the spreadsheet approach. In general, this model provides the ability to estimate the effects of crew reductions, emergencies, and modifications to the crew structure on the workload of ship personnel. Thus, this output illustrates how models can help establish a technical basis for manning decisions.

Analysis: This model has three primary limits. First, it includes no human performance data to illustrate how task demands might overload ship personnel. Failing to consider human performance limitations restricts its ability to predict task demands accurately. Specifically, the time and personnel allocated to "Docking" may require more detailed modeling, especially if significant changes have been introduced to reduce manning requirements. Second, the modeling tool supports only a description of tasks as a linear sequence of events, all of which have known lengths. The tool provides the ability to model tasks that begin or end at the conclusion of others, but it contains no provisions for modeling tasks with probabilistic lengths. For example, the actual time required for "unplanned maintenance" may vary dramatically, but the model may not reflect this variation. Therefore, when task varies dramatically in length, or a sequence of tasks cannot be predicted exactly, then the model may miscalculate the workload of ship personnel. Third, the model provides no explicit link between task demands and ship safety, such as probability of collision.

Recommendations: Despite the limitations inherent in the project management tool, and the limits of the data incorporated into the model, this model illustrates a potentially valuable technique for estimating the manning requirements of a ship. More specifically, this type of model could enable users to determine whether proposed manning configurations would enable ship personnel to attend shipboard tasks, while meeting workhour limits. Some of the issues that this model can resolve include: the effect on workload of shifting task responsibilities to different crew members, the effect on workload of reductions in crew size, the effect of automation and labor saving devices on workload, and the effect of voyage profiles on workload. In each of these instances workload is defined by the number of hours worked by ship personnel.

With a preliminary model including almost 400 tasks, this approach illustrates the magnitude of the effort, even when describing the system with fairly broad tasks like "Docking" and "Unplanned Maintenance". A more complete model would require a more detailed task description, requiring a significant data collection effort. Furthermore, including human performance limitations would require a much greater effort.

4.7 CONCLUSIONS

Examining these six models illustrates how limits associated with tools and content, as well as the complexity and diversity of issues involved with manning requirements, make it difficult to use any of these six models as a technical basis for all ship manning decisions. While no one model is comprehensive enough to examine all issues regarding manning decisions, many of these models can examine specific issues, such as the ability of a manning structure to meet workhour requirements. Thus, this analysis shows the feasibility of using models to address specific manning issues. Table 2 shows the capabilities of the six models. Those models with more black

Table 2. A summary of the model content for each of the six models.

Models / Tool Characteristics	Operational Sequence Diagrams	Petri Nets	Integrated System Reliability	Maritime Administration	National Research Council	Argent Marine
VARIABLE REPRESENTATION						
Continuous			●			
Discrete	●	●	●	●	●	●
Combined			●			
TASK ORDERING						
Concurrency	●	●	●	●		●
Priority		●	●			
Deterministic	●	●	●	●	●	●
Contingent	●	●	●			●
Probabilistic			●			
EVENT TIMING						
Deterministic	○	●	●	○	●	●
Contingent		●	●			●
Probabilistic			●			
RESOURCE REQUIREMENTS						
Deterministic	●	●	●	●	●	●
Contingent		●	●			●
Probabilistic			●			
RELATION TO SYSTEM						
Task sequence	●	●	●	●	●	●
Open loop			○			
Closed loop						

● Tool accommodates indicated dimension.

○ Tool marginally accommodates indicated dimension.

Table 3. A summary of the content for each of six maritime models.

Models / Attributes of Model Content	Operational Sequence Diagrams	Petri Nets	Integrated System Reliability	Maritime Administra-tion	National Research Council	Argent Marine
REPRESENTATION OF HUMAN						
Human performance		●	●			
Decision making	●	○	○			
Human reliability	○		●			
Crew interaction	●		●			
REPRESENTATION OF SYSTEM						
Broad scope			○	○	●	●
Detailed representation	●	●	●	○		
System performance	○	●	○			

● Model accommodates indicated dimension.

○ Model marginally accommodates indicated dimension.

dots represent more flexible tools, while those with fewer dots have fewer capabilities. While models with few capabilities may not be able to answer all manning issues, models with many capabilities may not be practical to develop because of difficulties collecting data and manipulating the modeling tool. This table shows that the tool developed by Siegel and Wolf has the greatest capability. However, this tool requires considerable expertise to operate, increasing the resources needed for development. Table 3 summarizes the content of the six models. The categories included in this table represent some of the characteristics that differentiate models as discussed in Appendix B. In particular, this table shows how each model represents the overall ship system, as well as ship personnel. Some models (Operational Sequence Diagrams, Petri Nets, and Integrated System Reliability) contain a detailed representation of ship personnel, while others (Maritime Administration, National Research Council, and Argent Marine) do not represent the factors that govern human performance. While these models fail to represent explicitly factors that govern human performance, the National Research Council and the Argent

Marine models describe the entire ship over an entire voyage. While no model provides a comprehensive representation of the ship system, each model has capabilities that show that manning model development represents a feasible method of estimating safe crew size.

Models like that developed by the NRC (which provided quantitative estimates of number of man-hours required for a voyage) illustrate some of the limitations of both the modeling tool and the model content. The modeling tool provided only summary estimates of personnel requirements with no consideration of peak workload or system reliability. In addition, the model's content lacked any human performance data that might have illustrated the interaction between personnel and automation, or the effect of fatigue on ship safety. Other models, such as the OSD approach of Kurke (1961) incorporated human performance considerations, but the content of the model included only a very limited scope of shipboard activities.

The limits of modeling techniques and model content, illustrated by the models in this chapter, reveal some of the problems with current manning models. However, the models discussed in this chapter were designed for specific functions and in most cases they were successful, illustrating the potential of developing a technical basis for ship manning. Although no one model provides a ready solution to the issues affecting ship manning, together they illustrate the feasibility of developing manning models to estimate safe manning requirements. While the individual manning models discussed in this chapter suggest that modeling can help support a technical basis for ship manning, the complexity of the ship system and the variety of issues makes developing a comprehensive manning model difficult. For example, even though the Argent Marine model contains nearly 400 individual tasks, it applies only to a specific ship on a specific voyage; furthermore it contains no human performance information. Likewise, the NRC and Maritime Administration manning models represent substantial efforts; however, they do not include a wide range of factors. A comprehensive approach to manning should include not only the information in the NRC, Argent Marine, and Maritime Administration models, but also more information concerning how automation alters the mental and physical workload. For example, the Argent Marine and NRC manning models do not explicitly account for the cognitive demands associated with varying automation and manning levels. These models implicitly represent cognitive demands based on expert judgment of time and skill requirements of shipboard functions. The operational sequence diagrams of Kurke (1961) illustrate how cognitive demands might be included in a model explicitly. Likewise, the model of maintenance requirements by Madni, Chu, Purcell, and Brenner (1984) shows the significant effort expended on estimating the cognitive demands associated with a single maintenance task. Thus, development of a technical basis for manning decisions requires either an extremely complex model, or a series of models that address specific manning issues. The approach adopted by Siegel and Wolf (1974), incorporated the broad scope of the NRC model, together with significant information concerning changes in human performance, such as stress, fatigue, and skill. As such, it is an example of a workable complex model. On the other hand, several simpler models could be developed which might include such factors as: evaluation of emergency preparedness, effects of automation, and estimation of overall workhour demands. This latter approach is discussed in the following chapter.

Developing a comprehensive approach for estimating manning requirements must go beyond previous studies. Thus, it would require several years of effort. Much of this effort involves determining what factors need to be included in the model and then collecting data to model these factors. An economical approach to this challenge is to develop a relatively simple model that determines manning requirements based on workhour limits, then enhance and expand this model as factors are identified and data is collected. The efforts of Argent Marine and the NRC illustrate that a model based on workhour limitations is practical, and collecting the requisite data is feasible. Likewise, the studies by Kurke (1961) and Madni, Chu, Purcell, and Brenner (1984) show that detailed models, developed to address specific issues, are also feasible.

V. COMBINING MODELING TOOLS AND CONTENT TO ADDRESS MARITIME MANNING ISSUES

Based on the results of the preceding chapters, it is clear that the development of maritime manning models is feasible. The critical aspect of developing such models is clearly defining the scope of the model(s) and the data requirements. This chapter begins with a description of the general requirements for a manning model, followed by a description of a general model development process. This chapter describes a strategy that begins with a relatively simple model, followed three models that serve to elaborate specific aspects of shipboard function. In addition, this chapter provides a strategy to integrate the results of these efforts, and concludes with a discussion of model validity and a validation strategy.

5.1 REQUIREMENTS FOR MANNING MODEL DEVELOPMENT

Model development requires more than creating computer programs and entering data. Initially, it requires accurately defining the range of predictions to be made. It further requires a large amount of time devoted to collecting data from either event reports, published reports, or field and laboratory studies. Furthermore, model validation is required to ensure that the model accurately reflects the actual system. Validation may require valuable simulator time and further data collection in the field. Figure 1 summarizes the process of model development, beginning with a definition of the scope and goals associated with the model. Following the definition of the scope and goals, data collection, and coding takes place. After a model has been coded, it can be used to examine different shipboard scenarios, answering the questions that inspired its development. Shipboard scenarios represent a combination of manning, voyage profile and equipment configuration. This combination of parameters represents the input to the model. Figure 1 shows that each step of the development process involves validation. In many cases the actual coding of the model constitutes a relatively small part of the overall development process; definition of modeling goals, collection of data, and model validation absorb the majority of the effort. The time required to develop a manning model, particularly data collection and validation, depends upon coordination with other projects. For example, if ongoing research can be tailored to include data collection for a manning model, then substantially less time would be devoted to this stage of model development.

The development of a hypothetical manning model may help to illustrate this process. The first step in the modeling process begins with a clear statement of the scope of the intended model and the type of questions that the model will resolve. This initial step specifies the type of modeling tool and content needed to develop a useful model. The second step of the process involves collecting data. For a manning model to evaluate workhour requirements, data might come from expert opinion, observation of shipboard activities, and the output of simulator studies. Specifically, a functional analysis that specifies the time and expertise requirements of shipboard

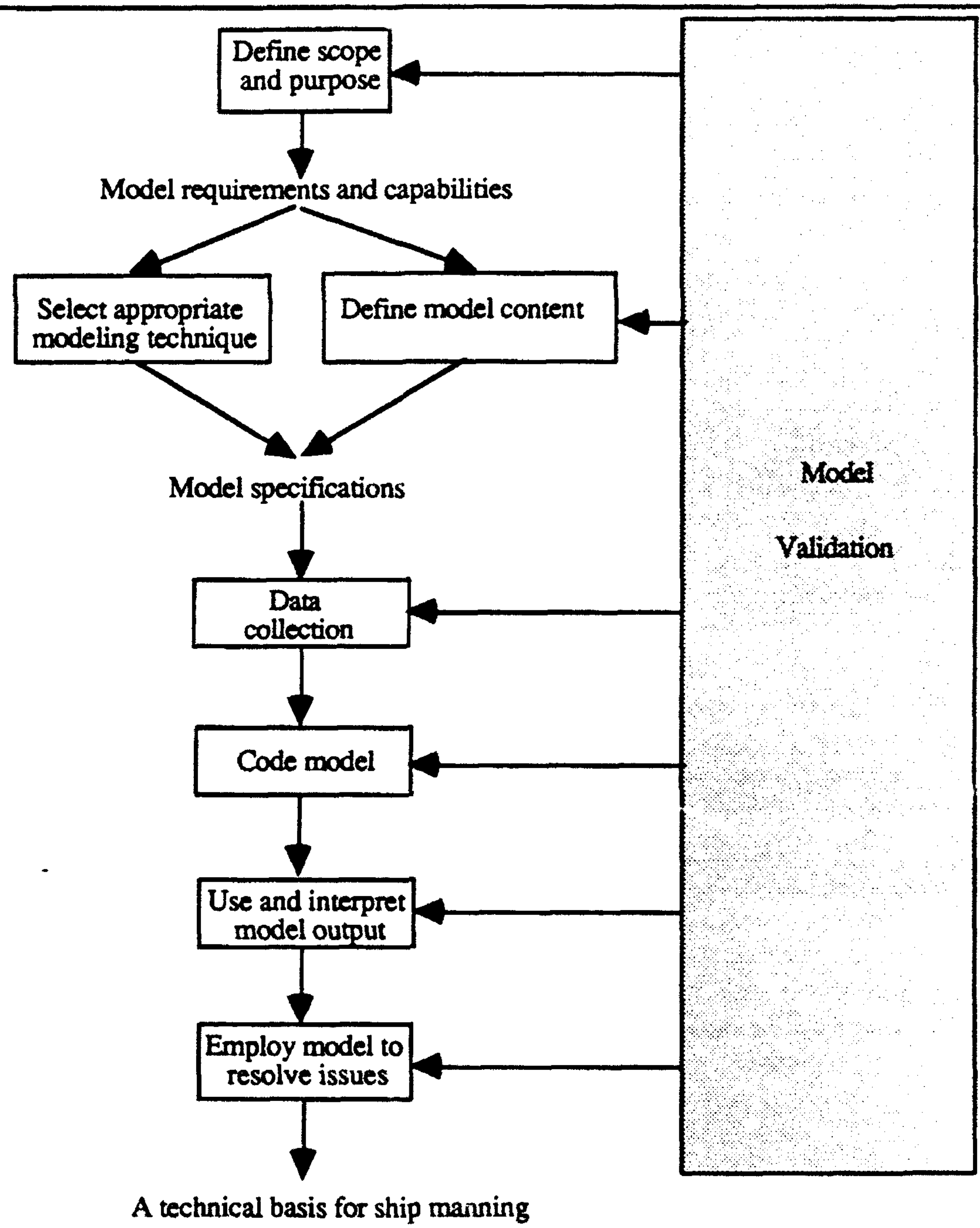


Figure 1. A flowchart showing the development and use of a model. The boxed text represents activities and un-boxed text represents information or output of the activities.

tasks could provide the basis for a manning model. Data are then translated into a formal model. Model validation occurs throughout the model development process, ending with a comparison of the predicted behavior to the behavior of actual ships. Figure 1 shows how model validation spans the development process. Model validation should address all steps of the development process. After successful validation, the model then provides information to answer questions identified in the first step of the model development process, such as how modification of the manning structure influences the ability of the crew to respond to emergencies.

5.2 SELECTING APPROPRIATE MANNING MODELS

As previously mentioned, specification of the scope and goal of the manning model is a crucial element of the development process. Because of the complexity of any large system, any one model will accurately portray only a small aspect of it. In addition, the resources required to construct even a limited model make development an expensive proposition. Therefore, it is important that the modeling technique and content support a model that accurately reflects aspects of the system that are of greatest interest. No single model will capture all the detail concerned with ship manning. Therefore, models must be selected to provide information to answer specific questions. For example, it may be impractical to develop a single manning model that provides a measure of the change in ship "safety" by including all the effects of reducing a ship's crew by 10 members. Instead, it may be more efficient to develop a model that examines how such a change in manning will affect emergency response or work distribution.

Figure 2 summarizes some factors that influence manning model development. This figure shows how model content and tools influence model accuracy and costs. Models that include more detail, represent the system more inclusively, and portray the system more accurately, require more resources to construct and maintain. Likewise, models that provide a more general description of a system (one that might generalize to a wide variety of similar systems) also require greater resources to construct. For example, creating a manning model that accurately reflects the behavior of a ship and its crew over all phases of a journey and during emergency situations, together with an ability to provide this information for a wide variety of ships (tankers, cargo ships, and cruise ships) would require enormous resources. Therefore, creating a manning model requires balancing the resources committed to the project with the desired detail, validity, and generality of the model. The overall process of development of a manning model involves combining tool(s) and content (task requirements, manning level, crew composition, and performance shaping factors) to produce a model that predicts ship safety. Performance shaping factors include the relationships between such factors as fatigue, stress, expertise, and human performance. Figure 2 provides an overview of how performance shaping factors and other model components combine.

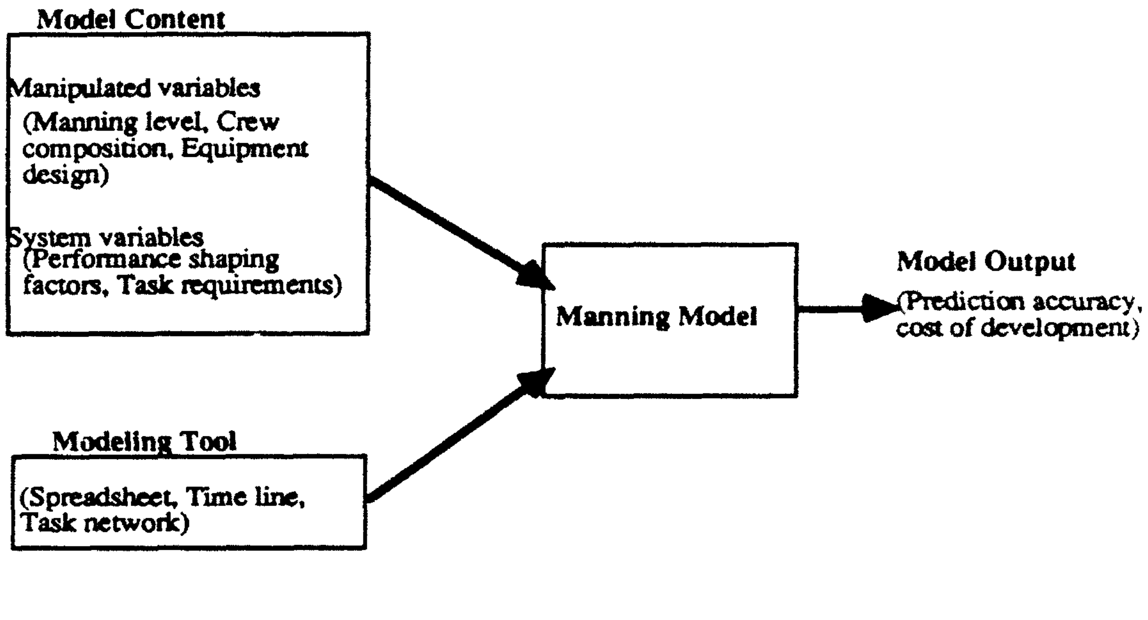


Figure 2. An illustration of how the model content and the type of modeling tool lead to development costs and manning model accuracy.

5.3 MATCHING MODELING TOOLS AND CONTENT TO ADDRESS MANNING ISSUES: A STRATEGY FOR MANNING MODEL DEVELOPMENT

The review of models applied to the maritime industry, presented in Chapter IV illustrates a great diversity of tools and content. Many of these applications successfully solved specific problems related to the issues involved with manning; however, none provide a comprehensive method of estimating manning requirements. Therefore, this review shows that it may not be possible to develop a comprehensive manning model. Instead the wide variety of issues related to safe manning may require a number of specialized models. However, as the work of Siegel and Wolf (1974) illustrates, large complicated models can produce estimates of the reliability of the entire ship system, taking into account a wide variety of factors. The development of such a model would require extensive resources. Therefore, this chapter describes the requirements involved in developing several separate models that each resolve specific issues.

This modeling effort should start with a model that shows how changes in the manning structure affect the ability of ship personnel to meet statutory workhour limits. Developing a functional model to evaluate workhour requirements would provide a foundation, upon which the Coast Guard could build more detailed models. This basic model could then be elaborated upon with more specific models that indicate the ability of the proposed crew to accommodate factors such as emergencies, maintenance, and port operations. The following section outlines a strategy for developing a functional manning model, followed by illustrations of models required to expand the scope and verify assumptions. For each of these models an estimate of the development time is included. This estimate is meant to illustrate the relative investment required for each model, and actual development time would depend on a more precise description of the scope and mission of the model. Thus, the following sections include: (1) the ability of ship personnel to meet workhour requirements specified by regulations, (2) the ability of ship personnel to manage emergency situations, (3) the effect of crew reductions, facilitated by automation, on the navigation performance, and (4) the ability of a reduced crew to meet demands of scheduled and unplanned maintenance. These issues were selected because they are representative of the types of issues that manning models should resolve. In addition, these issues require several modeling techniques, so the description of hypothetical models used to address these four issues illustrate the importance of matching the capabilities of the manning model to the issue. The following pages illustrate how the descriptions of modeling tools and model content in Appendices A and B help specify models most appropriate for each of these issues. The three issues, represented in the models used to augment the workhour requirements model, have been selected as representative of the types of questions that should be answered by a manning model and the modeling approaches that should be taken. The workhour requirements model and the three models that build upon it are not meant to represent an inclusive set of manning models. Instead, they represent an approach that will support a technical basis for manning decisions. This chapter concludes with a strategy for integrating and validating these and other models.

5.3.1 Workhour requirements evaluation to establish feasibility of proposed manning levels, using a timeline/project management approach

Problem statement: While increased automation and economic pressures act to decrease crew sizes, the number of hours worked by a ship's crew must meet regulatory standards. The Oil Pollution Act of 1990 limits the mariner's workday to 12 hours by limiting the number of hours worked to 36 every 72, and imposing a maximum workday of 15 hours. To receive a license and operate a ship, shipping companies must demonstrate the ability to meet these guidelines. An accurate estimate of the number of hours of work demanded of the crew requires a thorough understanding of the task demands associated with various technological enhancements, cargo types, and voyage scenarios. Proposals for flexible manning structures further complicate this problem, by making it possible to substitute various crew members for various jobs. The complexity involved with estimating a manning level that meets the workhour regulations makes it difficult to know if a proposed manning level meets regulatory standards.

Matching problem characteristics to a modeling tool: The primary characteristic of this problem involves understanding how time and resource demands of shipboard functions affect the workhours of crew members. In addition, the problem requires a preservation of the sequential relationship between tasks, so that no crew member is assigned to two tasks simultaneously. To address these issues, shipboard functions can be considered as a predetermined task sequence, with predetermined task initiations and durations. In addition, the resource demands (number of crew members and expertise requirements) for each function can be assumed known and constant. A "Project Management" tool accommodates these problem characteristics by combining characteristics of the timeline approach with some characteristics of a task network approach.

Matching problem characteristics to model content: Since this problem involves assessing the match between workhour requirements and crew composition, the model needs to contain only information regarding time required to execute functions. Therefore, the model need not contain detailed information related to human performance limitations or decision making abilities. Likewise, the model would contain no information concerning the performance of the ship, since the model assumes that a crew large enough to meet the workhour requirements of the shipboard functions will ensure ship safety. Since the purpose of the model is to provide a relatively simple evaluation of workhour regulations, the level of detail used to specify task durations would include fractions of an hour rather than fractions of a minute. The data requirements for this model could be met through direct observations, questionnaires and expert judgment.

Model input: The input to the model would be a table of tasks that specifies task durations, expertise requirements, number of crew required. This table could be supplied by ship owners, based on a template developed by the Coast Guard.

Resource requirements:
Data collection-- The resource requirement for data collection depends on whether it could be supported by other projects or provided by shipping companies. By tailoring a generic job/task analysis to include data in a format consistent with this modeling effort, the data could be easily collected. This job/task analysis could be used to develop a catalog of tasks that includes their durations, skill and knowledge requirements, and crew coordination needs.

Expertise-- Model development involves using a relatively simple approach, a project management tool. Therefore, no specific experience with the modeling tool would be required. However, the potential complexity involved in manipulating the data would require some experience in modeling. After the model has been configured relatively little expertise would be needed to enter data and interpret the output of the model.

Development time-- Assuming data have been collected and organized, the model could be completed in as little as four to six months. The exact time would depend on the level of detail and variety of ships and voyage profiles the model should support.

Model output: The output of this model would include a timeline illustrating personnel work commitments, showing the hours worked each day for the duration of the voyage. The timeline output provides a convenient method of evaluating whether the crew satisfactorily meets the workhour limits. Likewise, this output will reveal instances where the crew may be underutilized and suggest how tasks might be reallocated to distribute work more evenly. Since a project management tool prevents simultaneous allocation of a single crew member to multiple tasks, the model guarantees that all simultaneous tasks have been allocated to different people. Likewise, the model would ensure that tasks are assigned to people with the requisite skills. For this model, a criterion for safe manning might be the ability of the proposed crew to execute all the shipboard functions without exceeding workhour limits. These workhour limits might be those set by regulatory mandates, or they might be based on fatigue and vigilance research.

Model generality: This modeling technique has the potential of being general enough to encompass a wide range of ships. Since the model is based on a functional analysis of the ship, many of the general functions will remain constant across ships, only the durations, expertise requirements, and sequencing would change. By providing the shipping industry with a generic template for the functional analysis, which they can use to fill in the appropriate parameters, the model has the potential to apply to a wide range of ships and voyage scenarios.

Analysis: This type of model is a relatively simple representation of the ship system. To arrive at this simplified representation requires many assumptions. If these assumptions prove false then the validity of the model will suffer. For example, this model requires assumptions about the number of personnel, the expertise, and the time required to perform a function. If the actual function requires additional resources, then the model would underestimate the crew required for the ship. Similarly, the model contains no information concerning how human performance limits might affect ship safety; it assumes that the implicit human performance limits, represented by the regulations, ensure adequate performance. Under some circumstances such assumptions may not hold. For example, fatigue induced by long sea tours in rough weather may contribute more significantly to human performance, compared to working more than 15 hours in a single day. Likewise, some functions such as cargo handling may be especially susceptible to fatigue, and performance may suffer even with a 12 hour workday. This modeling approach also assumes both the ships' hardware and personnel operate reliably; however, people may become incapacitated and equipment may fail, and the manning estimates should include an analysis of the crew's ability to operate during small failures, as well as situations that develop into emergency situations. While this model has several limitations, many of these might be circumvented by testing critical assumptions with more specific manning models such as: *Automation impact analysis to consider cognitive demands associated with a highly automated bridge using a task network approach*, and evaluating the emergency response capabilities with *Emergency procedure evaluation using a timeline/project management approach.*

The decision to develop any one of these specific models could be supported by using the model to evaluate the effect of various assumptions on model output. Sensitivity analysis is a tool that reveals the magnitude of the effect of various assumptions. It involves manipulating model parameters and evaluating their effect on model output. Small changes in model parameters that result in large changes in model output may warrant closer scrutiny into underlying assumptions. Parameters that have little effect on model output do not deserve further investigation because more exact estimates would not influence decisions. An example of sensitivity analysis with this model might consist of manipulating the workhours associated with several functions. If small changes in workhour estimates of some functions result in significantly different manning requirements then these functions deserve closer scrutiny, either through empirical analysis or through more detailed models. More specifically, subtle changes in the workhour requirements associated with a highly automated bridge may result in significant changes in the need for highly skilled personnel. This might indicate a need to develop a specific model that focuses on a detailed representation of bridge operations.

Probability of Success (95%): This type of model is very likely to succeed. Previous developments by NRC and Argent Marine illustrated the general feasibility of this approach. This model is more complicated than the NRC model because it accounts for the workload of individuals across time, which will reveal workload peaks ignored by a spreadsheet approach that averages workloads across individuals and time. This model is also more complicated than the Argent Marine model because it will provide a representation of the ship system applicable to a wide variety of ships and voyage profiles, compared to the Argent Marine model that was developed for a single ship serving a single route. While the proposed model is more complex than either of these models, the history of this type of manning model (one based on a functional analysis of the ship) suggests that the added complexity will not introduce major problems into the development process.

5.3.2 Emergency procedure evaluation using a timeline/project management approach

Problem statement: A major concern associated with reduced manning lies in the ability of ship personnel to execute emergency procedures effectively. Under normal conditions, a ship may operate with relatively few crew members; however, emergency situations may exceed personnel resources. Since these emergency situations include oil spill mitigation, fire fighting, lifeboat deployment, and critical equipment failure, failing to deploy the appropriate number of crew could have severe consequences for human safety and the environment. A manning model could evaluate the ability of ship crew to meet the demands of emergency procedures. The underlying assumptions of this model include: current emergency procedures encompass important emergency scenarios, ship personnel follow these procedures, and that these procedures adequately accommodate emergencies. If these assumptions prove false, then additional data collection (ship personnel interviews and observation) would be needed to preserve the validity of the model. For example, missing or inadequate emergency procedures would decrease model validity, and might require considerable additional analysis to determine how the crew might respond. For example, oil spill mitigation procedures may not exist, or may be insufficient to address all but small spills.

Matching problem characteristics to a modeling tool: If present emergency procedures adequately cover the emergency situations that a ship would likely encounter, the primary elements of this problem relate to coordinating resources and communication over time. It follows that a modeling tool must capture the dynamic changes in resource requirements, as well as the coordination of resources required to support concurrent activities. For example, the tool must include the capability to represent expertise, leadership, and physical capabilities required for each task in the emergency procedure. In addition, since emergency procedures specify a sequence of tasks and activities, the successful execution of these activities translates into satisfactory system performance. More precise measures of effectiveness would be developed through interviews with training and operations experts. For example, these interviews might reveal that delays in certain tasks are intolerable while delays in others do not hinder system performance significantly. Project management modeling tools can also be used in this respect. They can identify the "critical path" in a chain of events, upon which the overall task time depends. Assuming that personnel execute emergency procedures as a sequence of fixed length tasks, with a minimal chance of incorrect execution, then a relatively simple modeling approach could represent the system. With these assumptions a modeling tool needs only to consider the system as a "Task sequence", and not in an "Open loop" or "Closed loop" manner.

Table 1 in Chapter III illustrates the capabilities of various modeling techniques, and shows that a timeline approach adequately represents many of the characteristics of this problem. "Project Management" tools combine characteristics of the timeline approach with some of those of the task network approach to produce a modeling tool that encompasses all the characteristics involved in emergency procedure evaluation (given the previously mentioned simplifying assumptions). A task network approach could provide a more accurate representation by including the effects of probabilistic task times and event outcomes.

Matching problem characteristics to model content: Because this modeling effort assumes correct execution of the steps specified by emergency procedures, the model includes very little need to represent the cognitive or physical activity of the crew. Instead, the model must include a representation of physical capabilities, leadership, and expertise required to execute tasks in the emergency procedure. This would depend on assumptions of the capabilities of the ship personnel, and would not be explicitly represented in the model. The model would simply show that some tasks could be performed by some personnel and not others. The validity of these assumptions would be assessed through interviews with subject matter experts. In addition, the content of the model must include information regarding concurrent activities to evaluate the ability of the proposed crew to meet the resource requirements of all the tasks. For example, fire fighting may require a team of five people trained to operate a fire hose, as well as a dedicated licensed officer to communicate with the bridge. In addition, the emergency may require the simultaneous need for a licensed officer to coordinate evacuation of the area of the ship affected by the fire. Therefore, the model must include information on the supply of adequately trained personnel and the simultaneous demands on those personnel. The issues of crew size and experience become even more critical when considering the failure of critical equipment. This type of emergency may require not only an enlarged crew, but it may also require very specific types of expertise, such as knowledge of Morse code if normal communication channels fail.

Assuming that execution of emergency procedures preserves ship safety, then the representation of the system need not be detailed or inclusive. The level of detail included in the model depends on the specificity of procedures and would probably include general tasks, spanning fractions of an hour and not minutes. The inclusiveness of the model's scope depends on the aspects of the ship specifically mentioned in the procedure. An exception might include the manning of the bridge; even if the procedure fails to mention the bridge watch explicitly, the model should include these requirements. Therefore, nearly all the data required for this model can be extracted from existing emergency procedures.

Model input: The input to this model consists of a description of the proposed crew complement, consisting of their skills, training, leadership capacities, and physical capabilities pertinent to the emergency procedures. This information would be organized in the form of a table that could easily be entered into the model.

Resource requirements: Data collection--The resources required to collect data for this model depend on its scope and the ability to tie the development to other research projects; much of the data collection needed to develop the model might come directly from a study that evaluates emergency procedures. Since the data required consist of a detailed listing of tasks involved in emergency procedures, much of this information could be extracted directly from emergency procedure documentation. In cases where explicit documentation does not exist, interviews with masters and crew members could reveal the tasks involved in an emergency response and the expected expertise and time requirements. Additional information regarding the physical capabilities, leadership, and expertise required for emergency procedures could be gathered through interviews with subject matter experts and observations.

Expertise-- Since the modeling technique used for this model is relatively simple, after an initial prototype has been built, additional modification and use of the model would not require special knowledge of modeling.

Development time-- Assuming data have been collected, the development of the models would be relatively fast. Approximately one week would be required to develop and validate a model of each emergency procedure. Thus, a model representing a collection of 24 critical emergency procedures would require approximately 6 months to develop.

Model output: The output consists of timelines of activities that indicate the ability of the personnel to meet the demands associated with emergency procedures. For example, a project management tool will indicate if a person has been allocated to concurrent tasks. Likewise, the tool will provide summaries of the tasks allocated to individuals and to teams. These results show when the demands associated with emergency procedures exceed the capacity of the crew. In evaluating the output of this model, one possible criterion for safe manning could be the completion of emergency procedures without the need to overburden ship personnel by requiring them to complete more than one task at the same time. Another criterion might be based on the time required to complete the procedures.

Model generality: Since this model depends on the task sequence specified by the emergency procedures, the model is as general as the emergency procedures. If certain procedures are specific to a ship, or class of ships, then the model will not apply outside their domain. For example, procedures governing oil spill mitigation will likely differ dramatically from those designed to stop gas leaks on LNG tankers. Likewise, as technology changes the types of emergencies and the procedures involved in their mitigation, the models will change. However, since many emergency procedures are common to many ships, the results of this model may apply to a variety of ships.

Analysis: One of the primary limitations of this model lies in the assumption that correct execution of the emergency procedures translates into satisfactory performance of the system. It is likely that many procedures would fail to accommodate emergencies and many procedures may not exist. Insofar as these deficiencies can be identified, procedures should be updated, providing a more valid representation of the system. Another limit of this model would lie in the assumption that ship personnel correctly execute emergency procedures, given the appropriate leadership, physical, and experience resources. This assumption ignores issues associated with decision making under stress, fault diagnosis, and physical and mental workload. Therefore, this model may fail to predict instances where the crew fails to execute emergency procedures correctly. Even with these limitations, this model would indicate whether a crew configuration meets the minimum requirements imposed by emergency procedures. Enhancing this model by using probabilistic data would increase its validity, but it would also require a more complicated modeling tool and much more extensive data collection.

Probability of Success (90%): While no other models have been developed to examine the effectiveness of ship personnel to execute emergency procedures, several factors suggest that this model could provide valuable data without excessive costs. For example, the data for this model should be readily available in the form of emergency procedures. These may be written procedures or implicit procedures developed through practice and understood by the captain and crew. In either case documenting them should require relatively little effort, and a model based on them will provide the Coast Guard with an estimate of manning requirements, as well as an estimate of procedure consistency and inclusiveness.

5.3.3 Automation impact analysis of cognitive demands associated with a highly automated bridge using a task network approach

Problem statement: Changes in technology onboard ships often promise to reduce manning requirements; however, simply reducing physical load may not reduce, and may even increase, mental workload. In these situations, a superficial examination of the situation may suggest an opportunity to reduce manning, but a more detailed investigation may show that workload has not dropped, but simply changed in nature. A wide variety of technological improvements have been introduced including light weight Kevlar lines, automatic tension adjustment on mooring lines, increasingly reliable propulsion systems, and a wide variety of navigation and collision avoidance aids. All of these forms of technological innovation may contribute to changes in manning requirements; thus, their effect on ship safety deserves investigation. The area that merits the closest scrutiny appears to be the bridge. Here mismatches between the demands imposed by ship control, the coordination of crew, and management of technological resources may exceed the capabilities of the bridge personnel. For example, increased centralization and automation may increase demands associated with monitoring, fault detection, fault diagnosis, and planning responsibilities. With the growing tendency towards centralizing ship management on the bridge, errors in judgment committed by bridge personnel are increasingly influential on ship safety. Therefore, a careful analysis of bridge operations might reveal the effect on ship safety of reduced or modified bridge manning that often accompanies technological development.

Matching problem characteristics to a modeling tool: Unlike the execution of emergency procedures, activity on the bridge involves more than a linear sequence of activities. Events on the bridge occur unexpectedly and may also depend on the previous activities of those on the bridge. For example, the appearance of another ship on the horizon or an alarm in the engine room occur probabilistically, while the point of closest approach exceeding safe limits may depend on the actions of those on the bridge. Likewise, use of technological and personnel resources may not always depend on normative procedures outlined in advance. In many cases the crew maintains the option to circumvent the automation and perform the task manually. Similarly, the officer on watch maintains the prerogative to call standby help to the bridge. In each of these instances the approach to using technological resources may contribute significantly to ship safety. To incorporate the relatively unstructured nature of tasks on the bridge requires a modeling tool capable of representing probabilistic events, and resource requirements, as well as a closed loop representation of the tasks in relation to the system. A closed loop representation facilitates the representation of the outcomes of tasks on future tasks, such as the effect of taking evasive action earlier rather than later. A task network approach satisfies most of these and other requirements. MicroSAINT, a task network simulation tool, provides additional features such as the ability to include both continuous variables (e.g., ship position and ship speed), as well as discrete variables (decision to call master to bridge).

Matching problem characteristics to model content: Cognitive limitations of the personnel contribute most significantly to the effectiveness of automation on the bridge. Therefore, a manning model must include a representation of the cognitive demands associated with a highly automated bridge. These psychological variables would include the number and frequency of tasks that require continuous monitoring, memory demands, interaction of fatigue and monotonous tasks, and detection of subtle signals. In addition, the model should include information relating to "Ship resource management", such as a representation of the ability of the captain and crew to manage the resources available on the bridge. For example, under what conditions does the watch-standing officer neglect to call for additional help? Likewise, the model should indicate the conditions that lead bridge personnel to fail to use available technology, or rely on technology beyond its capabilities. Even a simple model that represents the distribution of tasks among bridge resources could provide valuable insight into situations that threaten to overload the crew. The relatively focused scope of this model provides the ability to examine tasks imposed on bridge personnel at a fairly fine level of detail, which might include tasks lasting fractions of a minute, as opposed to fractions of an hour.

Existing literature will provide a source for some of this information; however, little research has addressed human reliance on automation (Lee and Moray, in press; Sheridan and Hennessy, 1984). Other sources for this information include detailed observation of actual operations, subjective ratings of task frequency and duration, and a task analysis based on the design specifications and task situation. Questions that require personnel to specify the percentage of time spent on different tasks, or the relative frequencies of tasks provide data that have been successfully used to support an analysis of the effect of automation (McCallum and Underwood, 1991). The observations may include quite informal, qualitative descriptions of behavior, as well as video tape recordings subjected to a more quantitative analysis. In some cases asking people to "think aloud" and analyzing the resulting verbal protocol can provide valuable insight into the mental components of tasks.

Model Input: The input to this model would consist of the proposed bridge personnel and their skills and training relating to the automation in the bridge. A more precise model might include general responsibilities, functions, and tasks assigned to the various personnel. This information would be organized in a table that could easily be entered into the model. Applying this model to different kinds of automation might require adjustments to task times, or even complete revision of how some tasks are represented. These adjustments would require expertise in model construction.

Resource requirements: Data collection-- Like other modeling efforts, the resources required to collect data for this model depend on its scope and the ability to tie its development to other research projects. By linking this model to studies of bridge automation much of the data for this model could be collected very economically. This is particularly important for this model because human interaction with highly automated systems is poorly understood and little theory exists to predict how automation might affect the workload of bridge personnel. Therefore, before a model can be built, substantial resources must be devoted to understanding the strategies operators adopt when confronted with a highly automated bridge. For example, technology may provide many labor saving devices, but personnel may ignore the features because of inadequate training, or lack of trust in the technology. The type of data that this model requires is a detailed cognitive task analysis that reveals the mental processes and strategies involved in the operation of a highly automated bridge. Thus, data collection must go beyond a simple description of physical activity, it must describe how personnel monitor ship status, interpret information, resolve ambiguities, and make decisions. These requirements make data collection for this model especially challenging.

Expertise-- Since model development involves creating a task network model, a substantial amount of modeling expertise will be required. In addition, the more complex model parameters, combined with a great variety of output data, require the model user to have substantial expertise.

Development time-- Assuming data have been collected, the development of the model should not require an excessive amount of time. However if data are required, then data collection will require a significant amount of time. The development time will be heavily dependent on the scope of the model. For example, a model that examines a wide variety of voyage scenarios will take much longer to develop compared to a relatively limited model that only examines a few scenarios.

Model output: The output of this model consists of a record of ship performance (e.g., time taken to identify an approaching vessel, point of closest approach, and time taken to change course) with different scenarios (manning configurations, weather, and traffic conditions). These metrics of ship performance could reflect the need for additional manning. In addition, this model could produce a variety of estimates of the mental workload imposed on ship personnel. One estimate might consist of a measure such as the one developed by North and Riley (1988), which considers interactions between different types of tasks (Wickens, 1984). Alternatively workload might be estimated through a method developed by McCracken-Aldrich (1984), which estimates workload through expert judgment of workloads associated with various dimensions of subtasks. For this model, a criterion for safe manning might be an absence of high workload peaks or continuously high workload. One of the simplest metrics of workload is the time required to perform tasks divided by the time available to perform tasks. If this measure exceeds 0.7 then manning levels should be increased (Bittner and Morrissey, 1990). By estimating both mental workload and ship effectiveness (point of closest approach, time taken to identify an approaching vessel, and time taken to change course) the manning model could show how the demands on ship personnel relate to ship performance and safety.

Model generality: The purpose of this model is to provide a detailed examination of the mental demands imposed on the bridge personnel by specific types of automation. Since these demands depend on the specific type of automation, the model will be less general compared to other models, such as the *Emergency procedure evaluation using a timeline/project management approach*. In addition, automation often affords a wide variety of strategies, and the specific strategies adopted may significantly influence workload (Adams, Tenney, and Pew, 1991). One of the challenges in developing this model is to provide estimates of workload demands associated with a variety of possible strategies. Therefore, both the diversity of potential bridge automation and the diversity of strategies adopted by operators act to limit the generality of results.

Analysis: One of the greatest limits of this approach to the development of a manning model lies in the process of data collection and model validation. Very little work has addressed the issue of predicting human interaction with automation, at least in a domain as complicated as the bridge of a large ship. Since automation may fundamentally change the nature of tasks, in ways that remain unanticipated before its implementation, observing the operator interact with the actual automated system is a crucial step in the development of a manning model. Surveys, observations, and protocol analysis can describe the effect of automation on ship performance, and support both model development and validation. Since the model should accurately predict the effect of the automation, the post-installation analysis should match the predictions of the model. Any mismatch between the model and reality provides an indication of how the model might be modified to reflect more accurately the effect of automation in a variety of scenarios. As discussed, this model assumes perfectly reliable automation and does not address the possibility of the failure of critical equipment. To include this important issue this model requires a link to the analysis of emergency procedures that address failures of critical equipment, discussed previously.

Probability of Success (65%): Models of this type have not been developed for the maritime industry, and analyzing and modeling human interaction with automation is still poorly understood in other domains as well. Therefore, little precedent exists to predict the success of this modeling effort. Even though it may be difficult to develop a model of how automation may affect bridge operations, such a model would be quite valuable. Even a limited success could provide valuable insight into the feasibility of different manning configurations and design options.

5.3.4 Scheduled and unplanned maintenance evaluation using a task network approach

Problem statement: The safety of a ship depends not only on the ability to navigate, but also on the ability of the crew to maintain the ship. Improper maintenance may jeopardize all ship functions, including safe navigation. Therefore, a significant effort should address the resource requirements and effectiveness of maintenance. Changing crew configuration and size poses particular problems related to maintenance. For example, reductions in manning may concentrate the workload on few people. Several major factors may combine with changes in manning practices to influence maintenance effectiveness. These factors include the resource requirements of scheduled and unplanned maintenance, and the dependence on shoreside support. These factors influence maintenance capabilities of the engine room, the radio room, and the computerized equipment installed on the bridge.

Matching problem characteristics to a modeling tool: The primary characteristics of this problem include the probabilistic occurrence of unplanned maintenance activity, the deterministic occurrence of scheduled maintenance, and the technological and personnel resources available aboard the ship and on shore. In addition, the problem includes tracking the effects of postponed or neglected maintenance. For example, the model should include some measure of task priority and the consequences for delay because some activities (chipping and painting) may have little effect on ship performance, while others (boiler leaks) may critically hamper ship performance. Therefore, for a manning model to investigate maintenance effectiveness requires the ability to include resource requirements, probabilistic task occurrence, probabilistic task durations, task priority, and deterministic task times. Table 1 in Chapter III shows that the task network approach meets these requirements.

Matching problem characteristics to model content: Unlike the manning model that estimates the effect of bridge automation on crew performance, the content of this model does not require an extensive representation of how task demands interact with cognitive limits. Instead, this model includes simple information regarding the time required for certain repair and maintenance operations. In addition, this model requires estimates of the probabilities of various failures and the schedule of planned maintenance activities. The final data requirement of the model is a listing of the personnel (expertise and number) and hardware (tools, diagnostic aids, and parts) required to perform the maintenance or repair tasks. Maintenance logs could provide all of these data and observation of actual repair and maintenance practices could provide confirmation. A more elaborate model might include human error probabilities to reflect the potential for failing to repair the system properly, or the chance that a scheduled maintenance activity would result in the system failing due to a technician's mistake.

Model input: The input to this model would consist of a table of the maintenance tasks included in the model. Each entry in the table includes the frequency of occurrence, the time required to complete the task, the number and expertise of the crew required to complete the task, and a measure of the consequences of neglecting the task. In addition, the model would require a listing of the personnel available for maintenance work, the number of hours a day they are available for maintenance work, and their expertise.

Resource requirements: Data collection-- The data collection for this model would require relatively little time and resources, compared to other models. Much of this information could come directly from existing repair records and maintenance schedules. Because the majority of the data have already been recorded, model development would only require organizing this information in a format consistent with the needs of the model development. The data collection would simply involve cataloging the frequency of various planned and unplanned procedures, together with the expertise and time requirements that they involve.

Expertise-- Since model development involves creating a task network model, a substantial amount of modeling expertise will be required. In addition, using the model requires knowledge concerning how to incorporate the statistical information about the distribution of maintenance tasks into the model.

Development time-- Assuming data have been collected and organized, the model could be developed quite quickly. If the data have not been collected, additional time of up to 3 months would be required to collect and organize the data. Depending on the number of maintenance tasks identified for inclusion in the model, development time could range from one to six months.

Model output: The output of this model would consist of summary measures of hours spent on maintenance and repair-related activities. In addition, the model could track the time delay involved in completing a maintenance or repair activity due to a lack of personnel or technological resources. For instance, some activities might go uncompleted because of a need to complete other, more important, tasks. In other cases, maintenance tasks might wait until the ship reaches the nearest port so that a part can be delivered. Another output of the model might be an estimate of the effect on ship performance created by the delay of a repair. In an extreme case, the crew's inability to complete a repair may incapacitate a ship. Such cases merit special attention as they represent emergencies and the manning requirements for mitigating their effects require further examination. For this model, the criterion for safe manning might set by establishing a limit of the total number of hours equipment can go without repair. This criterion could be further refined by weighting this measure by the criticality of the equipment.

Model generality: Depending on the outcome of an analysis of scheduled and unplanned maintenance records, this model might apply to a wide range of ships. For instance, if many maintenance activities are shared by most ships, then a model could be developed that includes these activities. Parameters specifying the exact maintenance schedule, frequency of unplanned maintenance, and resources required for repair could adjust this model so that it could be used for a wide range of ships. On the other hand, an analysis of maintenance activities may show that some activities are specific to certain ships, so the general model would not provide accurate results for those ships and the general model might require augmentation to accommodate these special maintenance activities.

Analysis: The primary limitations of this model include the potential problems of collecting data that accurately reflect the future repair and maintenance requirements of a particular ship. More specifically, using fleet averages might underestimate the maintenance required on particular trade routes, such as the Gulf of Alaska. In addition, as ships age, maintenance requirements will likely accelerate, and historical data will underestimate the number of future failures. Adjusting the probabilities to reflect these factors, as well as performing a sensitivity analysis to examine the effect of changes of the probabilities on manning requirements, would provide a reasonable estimate of the effect of different manning configurations on ship maintenance.

Probability of Success (95%): Models of this type have not been developed for the maritime industry to predict manning needs. However, the relative simplicity of the data requirements, combined with the potential for producing a highly general model, makes it likely that the model will provide useful data regarding the ability of different configurations of ship personnel to maintain a ship.

5.4 CONCLUSIONS

The four models proposed to examine issues involved with establishing safe manning illustrate the diversity of approaches that might apply to establishing a manning policy. Each of these models differs in its purpose, the tool adopted for its construction, and its content. Because each model addresses specific issues they all have limits; no one model completely resolves all the issues involved with determining safe manning levels. Instead, each model provides a focused examination of specific issues.

Table 4 summarizes the capabilities of the modeling tools used for each of the four models, while Table 5 summarizes the content of the four models. In particular, this table shows how each model represents the overall ship system, as well as ship personnel. Some models (Emergency response procedures and Automation impact analysis) contain a more detailed representation of ship personnel, while others (Scheduled and planned maintenance and Workhour requirements) do not incorporate detailed representations of the human element. While no one model contains

Table 4. The relationship between modeling tool capabilities of modeling tools and tool characteristics.

Tool Characteristics / Models	Emergency Response Procedures	Automation Impact Analysis	Scheduled and Planned Maintenance	Workhour Requirements
VARIABLE REPRESENTATION				
Continuous		●		
Discrete	●	●	●	●
Combined		●		
TASK ORDERING				
Concurrency	●	●	●	●
Priority		●	●	
Deterministic	●		●	●
Contingent		●		●
Probabilistic		●	●	
EVENT TIMING				
Deterministic	●	●	○	●
Contingent		●		●
Probabilistic		●	●	
RESOURCE REQUIREMENTS				
Deterministic	●	●	●	●
Contingent		●		
Probabilistic		●	○	
RELATION TO SYSTEM				
Task sequence	●		●	●
Open loop		●	○	
Closed loop		○		

● Tool accommodates indicated dimension.

○ Tool marginally accommodates indicated dimension.

Table 5. A summary of the model content for each of the six models.

Models / Attributes of Model Content	Emergency Response Procedures	Automation Impact Analysis	Scheduled and Planned Maintenance	Workhour requirements
REPRESENTATION OF HUMAN				
Human performance	○	●		
Decision making		●		
Human reliability		○	○	
Crew interaction	●	○	○	○
REPRESENTATION OF SYSTEM				
Broad scope	○		○	●
Detailed representation	○	●	○	○
System performance		●	○	

● Model accommodates indicated dimension.

○ Model marginally accommodates indicated dimension.

black dots in every row, representing a comprehensive manning model, combining the four models provides a broad range of capabilities.

If several models are used to address the diversity of issues that affect manning, a fruitful application of these models requires an integration of their results; otherwise a manning proposal may meet the requirements associated with one issue (e.g., overall workhour limits) but fail to satisfy the needs of another (e.g., emergency preparedness). Specifically, using the first model to evaluate a staffing proposal based on workhour requirements may conclude that the ship has been adequately staffed; however, the second model may reveal the crew's inability to respond to an emergency situation. Therefore, decisions concerning manning levels must depend on all four models, as well as others that address such issues as the effect of fatigue on decision making, ability of a reduced crew to dock a ship safely, and the demands on the crew imposed by cargo handling.

5.5 A PLAN FOR MANNING MODEL DEVELOPMENT

The review of existing models developed for the maritime industry, combined with the description of the four example models, suggests two possible approaches to developing a technical basis for evaluating the issues involved in determining safe manning levels: distributed and integrated. The distributed approach follows from the approach taken with the four example models: develop a series of models to evaluate specific issues, and to test assumptions of other models. With the distributed approach the feasibility of a staffing proposal depends on its ability to meet the criteria of several models. For example, a series of models must demonstrate the ability of the proposed crew to cope with emergencies, navigate with advanced technology, perform maintenance, and meet workhour limitations.

The integrated approach resembles that of Siegel and Wolf (1974), and involves the development of a single model with the flexibility to address all relevant issues. A task network modeling tool such as MicroSAINT could support such a model. The flexibility of this modeling tool facilitates incorporating a wide range of relationships that might influence manning. The tool facilitates a wide variety of output data, such as number of workhours, cumulative and instantaneous measures of workload, and probabilities of event occurrence. Building an integrated model hierarchically (models containing submodels and sub-submodels), provides method of organizing the information contained in the separate models of the distributed approach. In the extreme, a large integrated model could integrate hardware reliability with human reliability to produce a measure of the overall ship reliability. This number would specify the likelihood that the ship avoids a major accident during the next journey. For instance, a reliability of 0.999 would indicate a 1/1000 chance that the ship would encounter a major accident. Unfortunately, the effect of human behavior on system reliability is poorly understood and research continues to explore how human reliability can be integrated with machine reliability. While an integrated has the potential for developing a single index of ship reliability, it requires substantial resources for its development. Excluding the problems involved with using overall ship reliability as a the single metric of ship safety, using a single model involves added complexity because it contains all the information and capability needed to address a wide variety of issues. This added complexity makes validation more difficult and hinders subsequent modifications to the model.

With either the all-inclusive model (integrated approach) or a series of simpler models (distributed approach), a general modeling strategy might help integrate what would otherwise become a piecemeal and disorganized model. Bittner and Morrissey (1990) propose a systematic approach to model development, which includes models with different levels of sophistication. This approach could form a part of a strategy for evaluating manning requirements. One aspect of this approach involves concentrating effort on poorly understood aspects of the model. Developing only a simple model for parts of a system and a more elaborate model for other parts minimizes model complexity and conserves modeling resources. For example, the evaluation of a manning proposal might include a model of the second by second activity of bridge personnel interaction with proposed automation; however, such a detailed examination would be inappropriate for the steward's department. Figure 3 provides an overview of such a process,

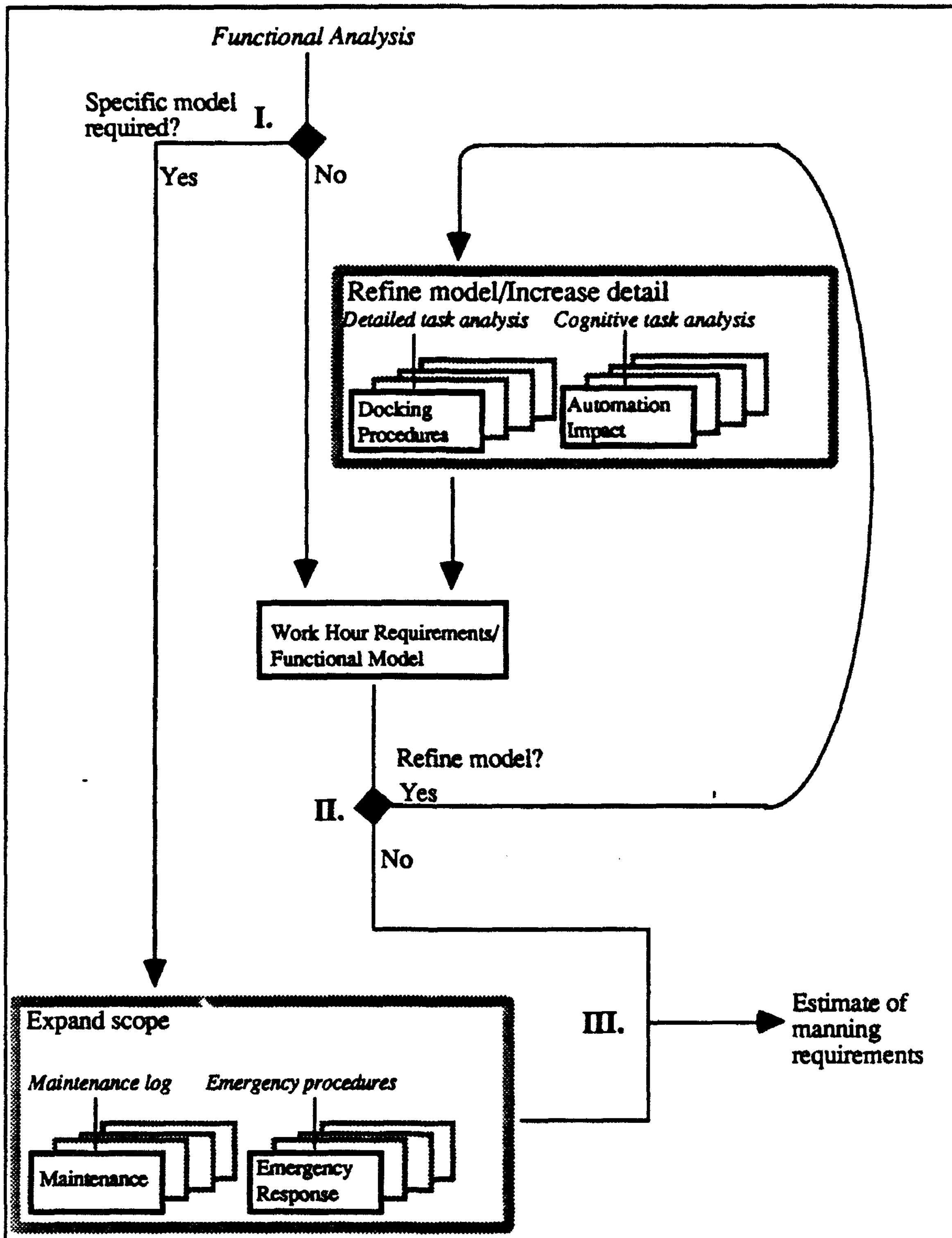


Figure 3. An illustration of a strategy for developing a technical basis for manning decisions (the stacked boxes represent alternative models and black diamonds represent decision points).

showing data sources (in italics), potential submodels (encircled in black rectangles), and major decisions involved in constructing a manning model (black diamonds). Major decisions include whether specific models are needed to expand the scope of analysis, and whether refinements to the assumptions of the original model are needed. Figure 3 shows that any approach to developing a manning model requires a functional analysis that specifies the capabilities a ship must meet. This analysis should include estimates of expertise and time demands. The NRC and Argent Marine models provide good examples of a functional analysis, and could be used as a first step in developing a more elaborate manning model. The functional analysis is shown as central to the modeling process in Figure 3.

The decision point identified by "I." addresses the need to expand the scope of analysis to give special consideration to specific functions to determine if ship personnel can meet specific task demands. For example, the functional analysis might reveal a wide variety of functions that require specific models, including maintenance and emergency capabilities. Specifically, a model of emergency response might specify manning requirements needed to address a variety of emergency scenarios. The results produced by this and other specific models would specify minimum manning requirements. For example, if an emergency scenario requires 12 people to accommodate, then the minimum manning level would be 12 people, even if the workhour requirements model identifies the minimum as 10 people.

A model of the global manning requirements, developed from the data in the functional analysis, shows whether the proposed manning structure meets statutory workhour limits. In addition, it shows what aspects of the model need refinements (represented by the decision point marked with "II."). Using sensitivity analysis, the most critical aspects of the model are identified and refined by collecting more data and developing more specific models. Figure 3 shows "Docking Procedures" and "Automation Impact" as two of many possible models that might help validate assumptions of the more global manning model. These more detailed models would show if the assumptions of the global model are accurate. For instance the "Automation Impact" model could indicate whether workhour assumptions, based on increased bridge automation, are justified. For example the results of this model (at least two people must man the bridge at all times) could specify the functional requirements used as input to the workhour requirements model.

The point marked with "III." represents the integration of the results of the workhour requirements model and the specific models developed to expand the scope of analysis. The integrated approach includes this step implicitly; with the distributed approach this step involves evaluating the output of several models. This evaluation process requires operationalizing the definition of "ship safety". For instance, with the workhour requirement model, safe manning depends on the ability to meet workhour restrictions. With a model of unplanned maintenance, having a greater probability of completing required unplanned maintenance tasks might constitute safe manning, when compared to a limit specified by the Coast Guard. Defining measures and associated criteria that constitute safe manning is a critical step in developing manning models to serve as a technical basis of ship manning.

5.6 VALIDATION: THE MATCH BETWEEN MODEL OUTPUT AND ACTUAL SYSTEM BEHAVIOR

The validity of a model depends on its ability to replicate the behavior of the actual system. Because any model represents a simplification or abstraction of the actual system, the behavior of the model will never be isomorphic with the behavior of the actual system. Therefore, any model will never achieve perfect validity; however, a model's behavior must replicate the behavior of the actual system closely enough to support decision making. Thus, while validity of a model will never be complete, it should be measured with respect to the purpose of the model and the decisions it supports. Model validity can be defined and measured in a number of ways. This chapter presents several definitions of validity from experimental psychology along with validation methods derived from systems engineering.

Researchers in experimental psychology have discussed a number definitions of validity (Cronbach and Gleser, 1965, Campbell and Fiske, 1959). More specifically, Cook, Campbell, and Peracchio (1991) mention four types of validity; Statistical Conclusion Validity, Internal Validity, Construct Validity, and External Validity. Applied to the evaluation of manning models Statistical Conclusion Validity refers to differences between the model and reality due to true differences as opposed to those due to random variation. Because of random variation, the behavior of actual ships will vary, therefore differences between the output of the model and the actual ship may simply reflect this random variation. Statistical tests help differentiate these random effects from consistent biases. Internal Validity refers to differences between the model and reality due to failures to incorporate important causal relationships. For example, failing to include human performance limitations (when the limits of human performance affect system performance) might lead to a model that fails to correspond to reality. Construct Validity refers to differences between the model and reality due to poor operationalization of theoretical constructs. An instance of poor construct validity might include defining "ship safety" (an unobservable theoretical construct) as the inverse of the number of hours worked by the crew. In this instance a wide range of other factors influencing "ship safety" would go unmeasured, possibly overestimating the true level of "ship safety". External Validity refers to differences between the model and reality due to qualitative differences in application domains. A model with high external validity could apply to a number of systems, over a variety of scenarios, besides those used for its development. In other words, External Validity refers to the generality of the model. For example, a manning model developed using task time data derived from a ship's crew that is well acquainted with each other may produce invalid results when applied to a crew who has never worked together. Each of the four types of validity limits the application of a model, and require consideration if the model is to support decision making adequately.

Naylor and Finger (1967) present three validation techniques that can address these four types of validity. These techniques include: technical examination of the model's underlying assumptions and logic (rationalist approach), empirical examination of model behavior compared to reality (empiricist approach), and a pragmatic approach, using the model to predict future system behavior. The rationalist approach to model validation depends on the premise that if one accepts the basic assumptions and the consequences of the logic incorporated in the model, then one must accept the model's validity. Unfortunately the complexity of many simulations may preclude this type of analysis. For example, many complex computer programs fail unexpectedly because the complexity of their logic makes it impossible to complete a thorough analysis. Chapters III and IV of this report provide a basis for an analytic examination of the underlying assumptions of a model as a first step in a rationalist approach to validation. Carefully examining the capabilities of the modeling tool and the content of the model may reveal inadequacies and limitations of the model even before its construction. A jackknife validation technique could be used to surmount this problem. It involves withholding some of the sampled data and using the rest of the sample to model it.

Although the rationalist approach may indicate potential limits of model validity, an empiricist approach provides a more direct estimation of model validity. Based on historical data, the behavior of the model can be compared to the behavior of the actual system. Mathematical approaches (analysis of variance, Chi-square, and regression techniques) help support the empirical analysis by comparing average values of system responses, distributions of responses, and correlation of patterns of responses with those from the model. One of the primary problems of the empiricist approach lies in collecting data for comparison. Data collection is so time consuming and expensive that seldom is enough data collected to support the development of a model, not to mention its validation. Even without the problems associated with data collection, the empiricist approach measures validity only in comparison to existing or past systems. Since the data from these systems formed the basis for the model, this method of validation does not test the ability of the model to predict the consequences of future changes. For example, the NRC (1991) report was validated by comparing its suggested manning level with that of the ships that provided the data for its construction, an approach that is tautological at best.

The final approach to validation might be considered pragmatic, with model validity depending on its ability to predict the future behavior of the system. This approach to model verification involves comparing the input/output transformations of the model to those of the system. For example, given a set of circumstances or parameter values, the model should produce behavior that predicts future system behavior. A wide variety of techniques might evaluate this correspondence, ranging from sophisticated mathematical approaches to techniques such as the "Turing test", where experts judge the behavior of the models as being representative of the actual system (Shannon, 1975). The appropriate technique depends on the level of analysis. In some instances, validation might compare sequences of behavior, in other instances it might involve comparing the average number of workhours. The appropriate level of analysis used in this comparison depends on the purpose of the model and the resources available for data collection. For instance, a model might be used to produce a distribution of possible task sequences; however, collecting enough data to determine the difference between the distribution of the model and that of reality may involve extraordinary resources. Thus, the greatest drawback of this validation approach, applied to the maritime industry, lies in the high cost associated with data collection (in terms of the time required to collect the data and in terms of the potential for catastrophic accidents). Therefore, experimenting with manning levels to validate a manning model must be conducted with great care.

While each approach to model validation has limitations, using a combination of the three approaches circumvents many of these problems. Because of the many factors that threaten model validity and the many constraints on the validation process, using a combination of three approaches will result in more accurate representation of the ship system than any single approach. The three approaches to model validation (rationalist, empiricist, and pragmatic) apply to any one of the four types of model validity (statistical conclusion validity, internal validity, construct validity, and external validity); however, different types of validity become important at different stages of model development. Figure 4 illustrates the process of creating a manning model and some of the threats to validity that are specific to the different steps involved in creating the model. The first step of model development (Define scope and purpose) provides the basis for validation and bounds the expectations of the model. Thus, improperly defined scope and purpose may lead to a model of little value because of its lack of generality (external validity). Similarly, selecting a technique and specifying the contents of the model may lead to an invalid model because these choices may ignore important causal relations (internal validity) and may fail to provide a firm basis for theoretical constructs such as mental workload (construct validity). For example, defining mental workload as an average time on task may ignore monitoring responsibilities and sudden peaks in task demands, underestimating the true mental workload. Coding the model represents another threat to model validity because the coding may contain typing errors, threatening the internal validity of the model because the model fails to contain correct relationships between tasks. Likewise, poor choice of values for input parameters and inadequate experimental design, along with failures in correctly interpreting model output lead to an invalid model (based on lack of statistical conclusion validity and construct validity). The rationalist, empiricist, and pragmatic approaches can test the validity of the model at each of step of model development.

A concrete example of this process will help illustrate the validation process. The test of the validity of the workhour requirements model begins with an evaluation of whether the functional analysis covers all the important functions of the ship system. This occurs when the functions included in the model are compared to the scope and purpose of the model specified by the Coast Guard, and the actual functional requirements of ship operation. This step of the validation process could be supported by providing those who review industry submissions with a list of required shipboard functions that must be included in a manning model. Any missing functions will threaten model validity. Furthermore, field inspection would illustrate whether the model includes all pertinent functions.

The second test of the validity of the workhour requirements model involves comparing the capabilities of modeling tool with the capabilities needed to represent the system. Appendix A of this report supports this step of the validation process by illustrating the capabilities and limitations of modeling tools.

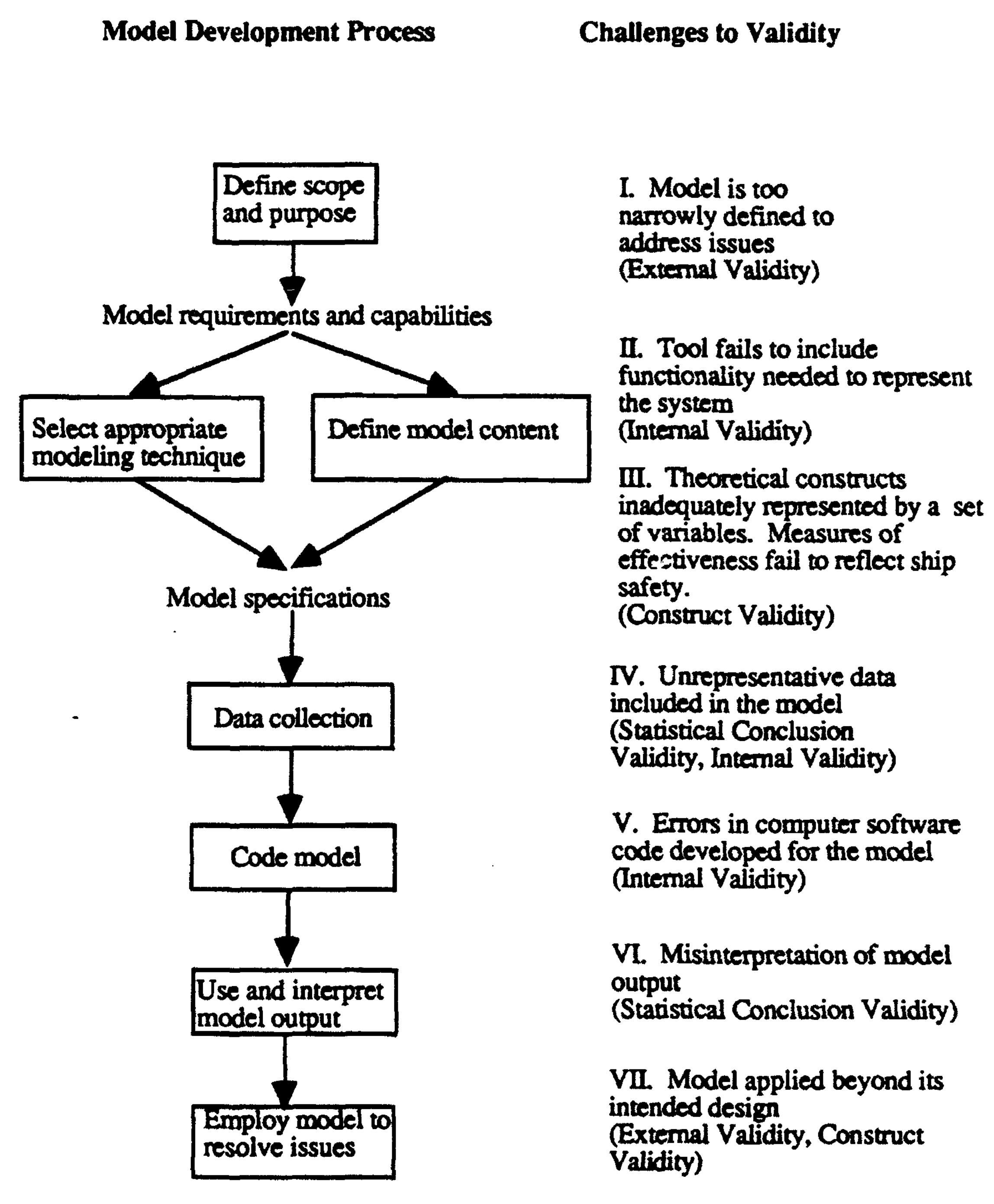


Figure 4. An illustration of the steps involved in creating a manning model and the possible challenges to validity.

The third step involves evaluating how well the measures of effectiveness used to evaluate model output relate to ship safety. With the workhour requirements model this involves estimating whether meeting workhour limitations indicates safe manning. This is a problematic step in the validation process and may involve additional model development, empirical testing, and subjective judgments of Coast Guard policy makers. For example, future research may show that certain functions are susceptible to fatigue and ship safety may be compromised by manning levels that force personnel to work more than 10 hours in 24. In terms of evaluating industry submissions, the Coast Guard could establish a standard format for model output. Using this standard format, industry submissions could easily be evaluated against specified criteria. For instance, an evaluator would check to see that the model produced estimates of the peak number of hours worked within a 24 hour period, and that this estimate does not exceed the statutory maximum of 15.

The fourth step in evaluating the validity of the workhour requirements model involves verifying the data included in the model. Specifically, this would involve verifying the assumed skills of the proposed crew, and the time and expertise requirements of each of the functions included in the model. Similar to testing whether the model includes the correct shipboard functions, a template could be prepared that shows the skills assumed of specific ship personnel, and skills and time required to serve different functions. Deviations from the standard would require the shipping company to demonstrate their validity. This step in the validation process could also extend to sea trials, where the list of functions and the associated personnel is validated against the actual activities of the crew. Mismatches between the proposed and actual crew assignments and the actual would threaten the validity of the model.

The fifth step in the validation process would involve checking to see that the software is correctly configured to combine task time and skill requirements with available ship personnel to produce estimates of workhour requirements. Since all the relationships in the workhour requirements model are deterministic (all task times, sequences, and skill are completely specified) no statistical analysis will be required, and so there is no need to verify the statistical validity of the interpretation of model output.

The sixth step in model validation verifies that the model is used for the purpose it was designed for. For example, if the model includes special provisions for manning reductions, based on a restricted trade route, then these provisions should restrict the use of the model.

To summarize, the following represents a list of the some of the steps needed to ensure that the workhour requirements manning model provides valid estimates of personnel needed to meet the functional requirements of a ship:

1) Verify functional listing

2) Verify capabilities of modeling tool

3) Compare output with validated measures of effectiveness

4) Verify time and skill expertise estimates associated with shipboard functions against standards and against actual practice during sea trials

5) Check software configuration to ensure accurate estimation of workhour requirements

6) Evaluate the match between the stated purpose of the model and the current application; establish whether special restrictions should be placed on the conclusions drawn from the model.

VI. CONCLUSIONS

This report addresses the need to develop a technical basis for ship manning. The goal of such a technical basis is to provide information that decision makers need to balance the effects of technological advances and economic pressures with ship safety. The models presented in this report illustrate the potential for estimating the effects of manning structures on ship safety. Although a variety of methods could serve this purpose (such as adjusting ship manning and evaluating the rate of subsequent accident reports), manning models offer a safer, more efficient, and more cost-effective approach.

This report examines modeling techniques in the maritime and other domains, such as aviation, military weapon systems, military command and control systems, and industrial process control. This review of models illustrates the diversity of potential approaches. Just as engineering models employed in ship design vary from structural, electrical and thermodynamic, to materials science, the variety of models that might apply to specific issues involved with safe manning vary dramatically. This diversity of potential models prompted the development of a framework for comparison of models based on a classification of modeling tools and content. This classification provides a basis to evaluate the ability of potential models to provide a valid basis for manning decisions.

Potential manning models range from simple spreadsheets, that add the number of hours that tasks aboard a ship demand of the crews' time, to much more complicated models that represent the effects of fatigue, workload, and types of automation on ship safety. To show how models might address specific manning issues, this report presents a strategy for model development, illustrated with four specific models. No one technique provides the capability to address all issues involved with ship manning; however, combining a workhour requirements model with other, more detailed, models will help establish a technical basis for ship manning. The report shows how several models might build upon the workhour requirements model to provide information for a technical basis for ship manning. This strategy supports two alternatives: (1) combine the content of sub-models into a single all-inclusive model, (2) evaluate the output of a number of specialized models. The most practical alternative would be the second approach (distributed), developing specialized models to address specific issues. Development of a single, all-inclusive model (integrated approach) introduces unnecessary complexity, and it would be more difficult to adjust to the specific needs. Therefore, the Coast Guard should begin developing a technical basis for ship manning by constructing a functional model to evaluate workhour requirements. Constructing such a model, based on the data and techniques of the NRC and Argent Marine, provides a cost-effective way of developing a basic manning model and evaluating the need for additional data collection and model development.

Central to the establishment of a technical basis for manning decisions are four components. The first component is a generic identification of tasks and duties involved onboard ships. Once general tasks and duties are identified, this template can aid ship owners in identifying the specific time and expertise requirements when they submit a manning proposal. The second component is the identification of a general profile of factors that should be included in manning models used to evaluate manning levels. These factors should include issues such as: human performance limits, effects of voyage scenarios, ability to accommodate unplanned maintenance, emergency response, and statutory mandates. These factors can be identified through functional and sensitivity analyses, performed using simple manning models that identify the critical functions a ship must perform. The third component essential to developing a technical basis for manning depends on the ability to validate the results of a manning model. This validation procedure should begin during the development of model specifications and continue through the evaluation of the effectiveness of the manning complement derived by the model. This procedure should include a comparison between the characteristics of the tool and content and those of the actual system. In addition, it should include a structured evaluation of ship performance through observations during sea trials, as well as examination of safety records. The final component of a technical basis for manning is a method to derive a baseline estimate of the manning requirement of any ship. The ***Workhour requirements evaluation to establish feasibility of proposed manning levels*** model could be the basis for this estimate. The baseline estimate could be refined with further analysis regarding the such factors as the level of automation, trade route and maintenance procedures.

Ultimately, developing a series of models to address safe ship manning will lead to a technical basis for manning requirements, similar to the technical basis that guides the construction and modification of ship hardware. Using a series of manning models to estimate ship manning will result in manning levels closer to "optimal" levels. Like engineering models that have led to "optimal" ship designs, revised manning levels may include economic advantages, as well as reductions in redundancy. Both optimal ship design and manning may suffice for any currently anticipated scenario; however, the reduced redundancy may leave ships ill-prepared for unusual emergencies. Blind reduction in manning could lead to accidents because the reductions did not take into account the such factors as the effects of crew workload, emergency needs, and maintenance requirements. The thrust of preparing a model is to accommodate various important factors (such as workload, emergency needs, and maintenance requirements). so that we do not make these horrendous mistakes.

Thus, the development of a technical basis for ship manning should begin by defining detailed specifications of a manning model to estimate workhour requirements, based on a functional analysis of the ship system. These specifications, developed in close coordination with the users of a manning model at Coast Guard headquarters, would identify its scope, flexibility, operation, and role in the regulatory process. The most practical method of generating these specifications would be through an iterative process, where manning model prototypes are developed and presented the Coast Guard staff. These prototypes could be quickly refined, based on Coast Guard critiques, and presented for further evaluation. The end product of this process would be a basic manning model that provides a first step towards a technical basis for manning levels. In addition, it will reveal the need for additional data collection and model development. The distributed approach to model development, illustrated in Chapter V would support subsequent elaboration of this model, providing a robust method of developing and evaluating manning proposals.

REFERENCES

Adams, M.J., Tenny, Y.J., & Pew, R.W. (1991). Strategic Workload and Cognitive Management of Advanced Multi-task Systems. Wright-Patterson Air Force Base, Ohio: Crew System Ergonomics Information Analysis Center.

Bainbridge, L. (1987). Ironies of automation. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), New technology and human error (pp. 271-283). New York: John Wiley & Sons.

Baron, S., & Corker, K. (1989). Engineering-based approaches to human performance modeling. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. Van Breda (Eds.), Applications of human performance models to system design (pp. 203-217). New York: Plenum Press.

Baron, S., Zacharias, G., Muralidharan, R., & Lancraft, R. (1980). "PROCRU": A model for analyzing flight crew procedures in approach to landing. Proceedings of Eight IFAC Work Congress, Tokyo, Japan, Vol. XV, pp. 71-76 (Also, NASA Report No CR152397).

Beville Engineering, Inc. (1992). Consultants in Human Factors Engineering for Process Control and Oil Refinery Operation. Dayton, Ohio: Beville Engineering, Inc.

Bittner, A.C. Jr., & Morrissey, S.J. (1990). Integrated performance and workload modeling for industrial and other system applications. Advances in Industrial Ergonomics and Safety II, pp. 857-864.

Bogner, S., Kibbe, M., Laine, R., & Hewitt, G. (1990). Directory of design support methods. Washington, D.C.: Department of Defense Human Factors Engineering Technical Group.

Bolin, S.F., Nicholson, N.R., & Smootz, E.R. (1990, February). Two howitzer crew drill models (ARI Research Note 90-13). Alexandria, VA: U.S. Army Research Institute for Behavioral and Social Sciences.

Booher, H.R., & Hewitt, G.M. (1990). Manprint tools and techniques. In H.R. Booher (Ed.), MANPRINT: An approach to systems integration (pp. 343-390). New York: Van Nostrand Reinhold.

Buck, J.R., & Maltas, K.L. (1979). Simulation of industrial man-machine systems. Ergonomics, 22(7), pp. 785-797.

Campbell, D.T. and Fiske, D.W. (1959). Convergent and discriminant validation by the Multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Chubb, G.P., & Hoyland, C.M. (1989). Systematic modeling of multioperator systems to evaluate design concepts. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. VanBreda (Eds.), Applications of Human Performance Models to System Design, pp. 295-311. New York: Plenum Press.

Cook, T.D., Campbell, D.T., & Peracchio, L. (1991). Quasi experimentation. In M.D. Dunnette & L.M. Hough (Eds.), Handbook of Industrial and Organizational Psychology, (pp 491-576). Palo Alto, CA: Consulting Psychologists, Inc.

Corker, K.M., Cramer, N.L., & Henry, E.H. (1990, June). Methodology for evaluation of automation impacts on tactical command and control (C^2) systems: Implementation

(Report No. AFHRL-TR-90-9). WPAFB, OH: Logistics and Human Factors Division, Air Force Human Resources Laboratory.

Cronbach, L.J., & Gleser, G. (1965). Psychological tests and personnel decisions. (2nd ed.). Urbana: University of Illinois Press.

Dawes, R.M., Faust, D., & Meehl, P.E. (1989, March). Clinical versus actuarial judgment. Science, 243, 1668-1674.

Dutton, J.M., & Starbuck, W.H. (1971, April). Computer simulation models of human behavior: A history of an intellectual technology. IEEE Transactions on Systems, Man, and Cybernetics, SMC-1(2), pp. 128-171.

Edwards, W. (1962b). Subjective probabilities inferred from decisions. Psychological Review, 69, 109-135.

Elkind, J.I., Card, S.K., Hochberg, J., & Huey, B.M. (1990). Human performance models for computer-aided engineering. New York: Academic Press.

Farooq, M.U., & Dominick, W.D. (1988). A survey of formal tools and models for developing user interfaces. International Journal of Man-Machine Studies, 29, pp. 479-496.

Froese, J. (1987). Current development in Federal Republic of Germany regarding crew reduction and bridge automation. Presented at the Ship Operations, Management and Economics International Symposium. pp. 9-1 - 9-6. Jersey City, NJ: Society of Naval Architects and Marine Engineers.

Gaffney, M.E. (1987). Reduced manning in the liner trades: Technological capabilities and organizational implications. Washington, DC: National Academy of Sciences Maritime Transportation Research Board.

Jones, P.M., & Mitchell, C.M. (1991). Human-machine cooperative interaction in the supervisory control of a complex dynamic system. In 1991 IEEE International Conference on Systems, Man, and Cybernetics, Vol. 2, pp. 1301-1306.

Joseph, J.E. (1987, September). US manning: Past, present and future. District 2 MEBA-AMO.

Kirwan, B. (1988). A comparitive evaluation of five human reliability assessment techniques. In B.A. Sayers (Ed.), Human Factors and Decision Making. Oxford: Elsevier. 87-104.

Knudsen, R.K., & Mathiesen, T. (1987, Operational safety and minimum manning. Presented at the Ship Operations, Management and Economics International Symposium. pp. 27-1 - 27-10. Jersey City, NJ: Society of Naval Architects and Marine Engineers.

Kurke, M.I. (1961). Operational sequence diagrams in system design. Human Factors, 66-76. Great Britain: Spottiswoode, Ballantyne & Company Ltd.

Lane, N.F Strieb, M.I., Glenn, F.A., & Wherry, R.J. (1981). The human operator simulator: An overview. In J. Moral, & K.F. Kraiss, (Eds.), Manned Systems Design: Methods, Equipment, and Applications. New York: Plenum Press.

Laughery, K.R., & Laughery, K.R. (1987). Analytic techniques for function analysis. In G. Salvendy (Ed.), Handbook of Human Factors. New York: John Wiley and Sons, pp. 329 - 354.

Laughery, R., Drews, C., & Archer, R. (1986). A micro SAINT simulation analyzing operator workload in a future attack helicopter. Proceedings NAECON. New York, New York: IEEE.

Lee, J. and Moray, N. (in press). Trust and the allocation of function in the control of automatic systems. Ergonomics.

Madni, A.M., Chu, Y., Purcell, D., & Brenner, M.A. (1984, November). Design for maintainability with modified petri nets (MPNs): Shipboard propulsion system application (Contract No. N00014-82-C-0683; Tech. Report No. PFTR-1125-84-11). Woodland Hills, CA: Perceptronics, Inc.

Malone, T.B., Baker, C.C., & Permenter, K.E. (Unpublished contractor's report). The role of human engineering in ship manning reduction. Carlow Associates, Inc.

McCallum, M.C., & Underwood, J.A. (1991). Impact of automation on command and control information processing. Prepared for Presentation to the 59th MORS Symposium.

McCracken, J., & Aldrich, T.B. (1984). Analysis of selected LHX mission functions. Technical Note ASI 479-024-84(B). CA: Anacapa Sciences, Inc.

McKaig, A.E. (1985). Approach to modeling the distribution of the message service times for the fire support team experiment. (Memorandum report BRL-MR-3461). US Army Ballistic Research Laboratory, Aberdeen Proving Ground, MD.

Melber, B.D., Berk, B., Olson, J., & Tunestam, B. (1983). An international comparison of manpower and staffing regulation and practice in commercial nuclear power plants, (Contract FIN #B2360). Division of Human Factors Safety, Office of Nuclear Reactor Regulations, U.S. Nuclear Regulatory Commission, (NUREG/CR-2953, PNL-4469). Seattle, WA: Battelle Human Affairs Research Centers.

Moray, N., Ferrell, W.R., & Rouse, W.B. (Eds.). (1990). Robotics, Control and Society: Essays in honor of Thomas B. Sheridan. New York: Taylor & Francis.

Moray, N., & Lee, J. (1991). Trust and allocation of function between human and machine. Presented at the ASME Winter Annual Meeting, Atlanta, Georgia. New York: The American Society of Mechanical Engineers.

Munro, H.P., Martin, F.W. & Roberts, M.C. (1968). How to use simulation techniques to determine the optimum manning levels for continuous process plants. In E. Edwards & F.P. Lees (Eds.), The Human Operator in Process Control. London: Taylor and Francis.

National Research Council (NRC). (1990). Crew Size and Maritime Safety. National Academy Press: Washington, D.C.

Naylor, T.H., & Finger, J.M. (1987, October). Verification of computer simulation models. Management Sciences, 14(2), B-92 - B-101.

North, R., & Riley, V. (1988). W/INDEX: A predictive model of operator workload. In G. MacMillan (Ed.), Human performance models. NATA AGARD Symposium, Orlando, FL.

Park, C.S., & Noh, Y.D. (1987). A port simulation model for bulk cargo operations. Simulation, 48(6), pp. 236-246. (Supported by the Sea Grant under contract NA81AA-D-0050 with the U.S. Dept. of Commerce.)

Parks, D.L. (1978). Current workload methods and emerging challenges. In N. Moray (Ed.), Mental Workload: Its Theory and Measurement. New York: Plenum Press.

Parks, D.L. and Boucek (1989). Workload prediction, diagnosis, and continuing challenges. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. Van Breda (Eds.), Applications of human performance models to system design (pp. 203-217). New York: Plenum Press.

Payne, J.W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. Organizational Behavior and Human Performance, 22, 17-44.

Payne, J.W., Bettman, J.R., and Johnson E.J. (1988). Adaptive strategy selection in decision making. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 534-552.

Perse, R.M., Baker, C.C., Malone, T.B. (1990). Simulation analysis of human task loading for US Navy surface combatant damage control organizations (Prime Contract No. N00024-85-D-4373). Naval Sea Systems Command.

Perse, R.M., Callahan, K.P., & Malone, T.B. (Unpublished contractor's report). ISMS human engineering development of a task model and analytical tool. Final Technical Report (Prime Contract No. N00024-85-D-4373). Prepared for Naval Sea Systems Command. Fairfax, VA: Carlow Associates, Inc.

Pew, R.W., Baron, S., Feehrer, C.E., Miller, D.C. (1977). Critical review and analysis of performance models applicable to man-machine system evaluation (Contract No. F44620-76-0029). Washington, D.C.: Air Force Office of Scientific Research.

Pew, R.W., & Baron, S. (1983). Perspectives on human performance modelling. Automatica, 19(6), pp. 663-676. Great Britain: Pergamon Press.

PRC. (1992). Mid-term sealift technology development program: Task 1.5 assessment of advanced manning techniques. (Interim projection of manning levels final report.)

Rasmussen, J. (1986). Information processing and human-machine interaction: An approach to cognitive engineering. New York: North-Holland.

Reason, J. (1990). Human error. New York: Cambridge University Press.

Rouse, W.B., Reid, R.E. (1981). An approach to analysis of human-computer interaction in ship control. Proceedings of the Sixth Ship Control Systems Symposium, Vol. 2 (pp. E1 3-1 - E1 3-6).

Seitz, Capt. F.C. Jr. (1981, October). Shipboard manning reduction - how few will do? Proceedings. pp. 50 - 54.

Shannon, R.E. (1975). Systems simulation: The art and science. Englewood Cliffs, NJ: Prentice-Hall.

Sheridan, T.B. (1987). Supervisory control. In G. Salvendy (Ed.), Handbook of human factors. New York: Wiley.

Sheridan, T.B., & Hennessy, R.T. (1984). Research and modeling of supervisory control behavior: Report of a workshop. Washington, DC: National Academy Press.

Siegel, A.I., & Wolf, J.J. (1969). Man-machine simulation models: Psychosocial and performance interaction. New York: Wiley-Interscience

Siegel, A.L., Wolf, J.J., & Lautman, M.R. (1974, November). A model for predicting integrated man-machine system reliability: Model logic and description (Technical Report No. N00024-72-C-1277. Wayne PA: Applied Psychologyical Services, Inc.

Stassen, H.G. (1989). On the modeling of manual control tasks. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. Van Breda (Eds.), Applications of human performance models to system design (pp. 107-122). New York: Plenum Press.

Strauss, H., & Stein, J. (1981). Report of the President's task force on aircraft crew complement. (Report PB81 2257160).

Swain, A.D., & Guttmann, H.E. (1983). A handbook of human reliability analysis with emphasis on nuclear power plant applications. (NUREG/CR-1278). Washington, DC: USNRC.

Underwood, J.A., McCallum, M.C., O'Herin, C.M., Penner, D.D., & Tyler, K.K. (1991). Process flow analysis and effects of automation data collection and analysis plan for III corps warfighter (Contract No. DE-AC06-76RLO 1830). Washington, DC: U.S. Department of Energy.

U.S. Maritime Administration Office of Research & Development. (1983, February). Crew rationalization study: ODS liner vessels. (Draft Research Memorandum). Washington, DC: Author.

Veldhuyzen, W., & Stassen, H.G. (1977). The internal model concept: an application to modeling human control of large ships. Human Factors, 19(4), pp. 367-380.

Wickens, C.D. (1984). Multiple resources in attention. In R. Parasuraman & R. Davies (Eds.), Varieties of attention. New York: Academic Press.

Wiener, E.L. (1985). Human factors of cockpit automation: A field study of flight crew transition (NASA Contractor Report CR-177333). Moffett Field, CA: NASA-Ames Research Center.

Williams, V.E. (1983, February). Crew rationalization study ODS liner vessels. Washington, DC: U.S. Maritime Administration, Office of Research and Development.

Woods, D.D., Pople, H.E., Jr., & Roth, E.M. (1990). The cognitive environment simulation as a tool for modeling human performance and reliability (NUREG/CR-5213 Vol. 2). Pittsburgh, PA.: Westinghouse Science and Technology Center.

Woods, D.D., Potter, S.S., Johannesen, L. and Holloway. (1991). Human interaction with intelligent systems: Trends, problems, new directions. Cognitive Systems Engineering Laboratory Report. Columbus, Ohio: Ohio State University.

Woods, D.D., Roth, E.M., Hanes, L.F. (1986). Models of cognitive behavior in nuclear power plant personnel. (NUREG/CR-4532 Vol. 2). Pittsburgh, PA.: Westinghouse Science and Technology Center.

APPENDIX A FEATURES OF MANNING MODELING TOOLS

This Appendix provides a classification of modeling tools to illustrate their various capabilities, followed by a comparison of the capabilities of these tools based on the classification. The comparison of different modeling tools illustrates how their capabilities and limitations influence the ability of a model to represent the ship system accurately. Carefully matching the capabilities of the modeling tool to the characteristics of the system represents the first step in model validation. The following classification of modeling tools provides a framework for evaluating the capabilities and limitations of modeling tools, as a basis for evaluating manning models.

A.1 DESCRIPTION OF TOOL FEATURES

A.1.1 Event /variable representation

Representing systems by using continuously changing variables, as compared to discrete events, constitutes a fundamental difference between modeling perspectives. In the context of ship manning models, discrete events might consist of specific tasks ship personnel must complete, while continuous variables might represent changes in the position and speed of the ship in response to control commands. Some techniques represent systems with only discrete events, while others use only continuous variables, and some use a combination of both discrete events and continuous variables. Thus, these different perspectives define three general classes of modeling tools: **Discrete**, **Continuous**, and **Combined**. The merit of these different approaches is not absolute, but depends on the nature of the system being modeled.

Figure 1 illustrates how these different modeling perspectives might represent the behavior preceding a collision of two ships. One representation, using only discrete events, illustrates communication, specific actions, and events. The second representation, using continuous variables, shows ship behavior as changes in the distance that separates them. The last representation shows how the combined discrete and continuous representation incorporates both discrete and continuous variables. A combined representation, might provide the best model of the collision of two ships; however, a discrete event representation might provide the best representation of other scenarios like position fixing.

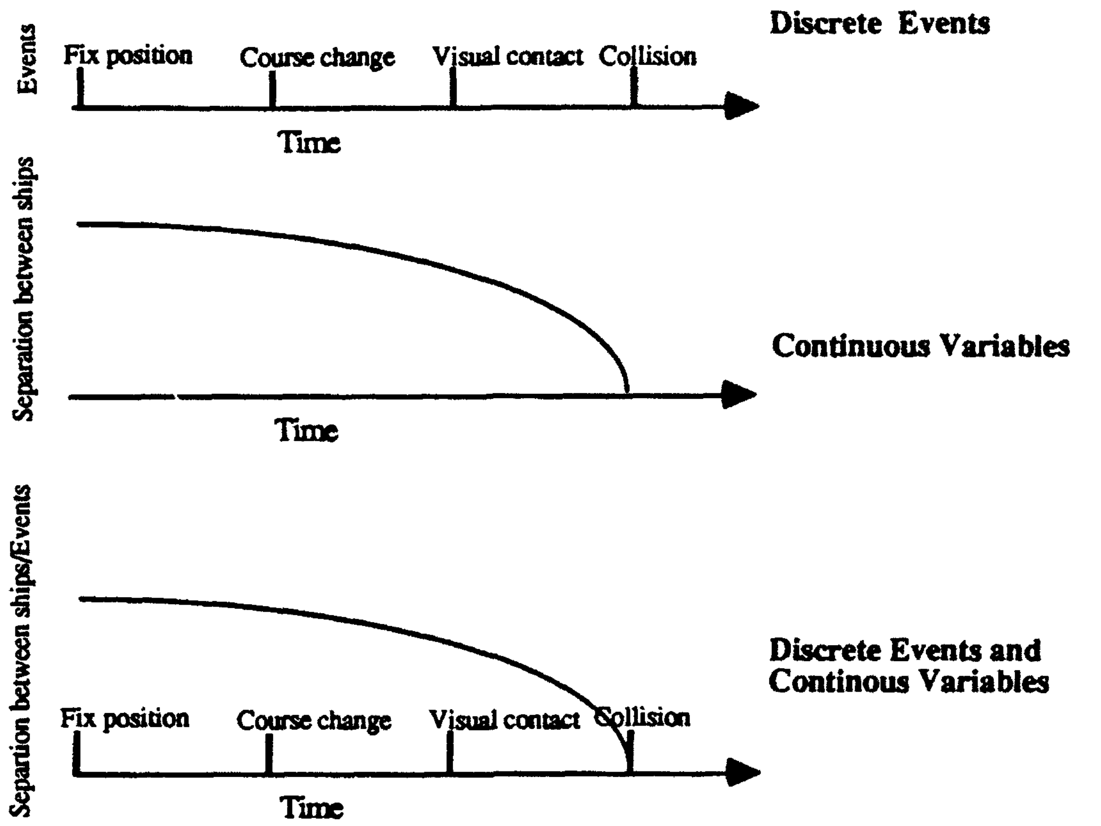


Figure 1. An illustration of how continuous variables, discrete events, and both discrete events and continuous variables would represent a collision of two ships.

A.1.2 Relationships between events/variables

Modeling tools incorporate various aspects of the system into a model by specifying **Deterministic, Contingent,** or **Probabilistic** relationships between the events/variables. **Deterministic** events have fixed start and stop times, and outcomes exist independent of either changes in system state or in the environment. For example, a manning model may specify **Deterministic** start and stop times of the position fixing task, beginning at 12:00 and ending at 12:06. With **Contingent** events, start and end times and outcomes depend on changes in system state, the environment, as well as previous events. With the task of position fixing, the captain's request might define the start time of the task that lasts 3 to 10 minutes, depending on the officer's fatigue. **Probabilistic** events do not depend upon other aspects of the system like Contingent events, but they do not have the known start and stop times of **Deterministic** events. Instead, **Probabilistic** events may have start and end times drawn from a distribution, and they may have indeterminate outcomes. For example, the mate may begin the position fixing task at any time from 11:55 to 12:05 and it may last 5 to 10 minutes, and he/she may not complete it accurately. Figure 2 illustrates differences between different types of representations of position fixing, showing the fixed time associated with **Deterministic** representation, the dependency associated with **Contingent** representation, and the time distribution associated with **Probabilistic** representation. The **Deterministic, Contingent,** or **Probabilistic** nature of events are not mutually exclusive and some modeling tools facilitate specification of events with a combination of all three.

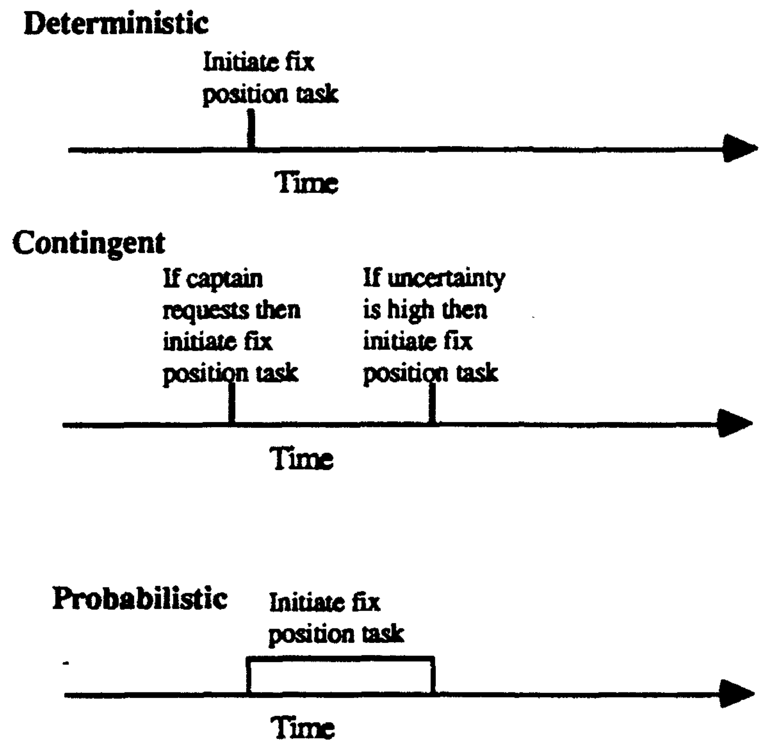


Figure 2. An illustration of the difference between deterministic, contingent, and probabilistic representation of the onset of different events.

A.1.3 Event ordering

The relationships between events specify the ordering of events in time and with respect to each other. Besides the **Deterministic, Contingent,** or **Probabilistic** nature of events, **Concurrency** and **Priority** can help describe how events occur over time. **Concurrency** specifies whether, and under what conditions, two or more events might occur at the same time. Imposing constraints on events, based on whether they can occur at the same time, facilitates a more accurate model of a system. For instance, with the tasks of position fixing and data logging, one model might account for the added mental strain (or the impossibility) of simultaneously executing both tasks, while another tool might ignore the cost of concurrently executing both tasks.

Incorporating the **Priority** of tasks into the model may play an important role in the ability of the tool to model the system accurately. If the model does not account for differences in urgency or importance between tasks, then it may not accurately reflect system performance. For example, tasks such as responding to a nearby ship, may be more important than fixing the ship's position, and the officer may delay fixing the ship's position to respond to the nearby ship. A model that fails to incorporate priority of tasks may incorrectly predict that the officer would fix the position of the ship first and then respond to the nearby ship.

A.1.4 Event timing

Some tools provide the ability to model the timing of individual events that occur in the system. A modeling tool may represent the completion time of individual tasks as either **Deterministic**, **Contingent**, or **Probabilistic**. As discussed previously, **Deterministic** event times imply that the start and end times are known exactly and the event occurs independently of system state. **Contingent** event times depend on personnel resources, system state, and the outcomes of previous tasks. **Probabilistic** event timing, on the other hand, results in events beginning and ending not at fixed times, but at times drawn from a distribution of possible times.

A.1.5 Resource requirements

Some modeling tools enable the model to include information about the resources required to complete individual tasks. Resources include both human and material requirements of the system. For example, for the task of plotting a course, the person plotting the course and the chart used for plotting might represent resources required for the task, without which the task remains uncompleted. As with the timing of events, the resource requirements of a task may be either **Deterministic**, **Contingent**, or **Probabilistic**. In the case of **Deterministic** resource requirements, the requirements of the task are known and invariant. They do not change over time or in response to changes in the system. The task of plotting a course, for example, requires both a person and a chart. In contrast, a modeling tool may represent tasks with a **Contingent** resource requirement, such as when the resource requirements depend on either system state or previous tasks. For example, ship handling in a narrow channel may require different skills or personnel, compared to ship handling on open seas. **Probabilistic** resource requirements might reflect variations in task demands, such as the occasional need for confirmation of a navigational calculation by other bridge personnel.

A.1.6 Event outcome

Just as some modeling tools allow a model to incorporate resource requirements, some tools provide a method of representing several possible outcomes for a single event. As with resource requirements, the outcome of the event may be either **Deterministic**, **Contingent**, or **Probabilistic**. **Deterministic** outcomes are completely specified. In other words, the outcome of the event remains the same every time it occurs. A modeling tool that incorporates **Contingent** outcomes provides a method of modeling events whose outcomes are not known in advance, but depend on system variables, previous tasks, or resource availability. "If...then" conditional statements can often represent **Contingent** outcomes, where the occurrence of an event or the value of a variable relates conditionally to other variables or events. Another outcome that a modeling tool might incorporate is **Probabilistic**. Like **Contingent** outcomes, **Probabilistic** outcomes represent situations where the outcomes of events are not completely specified during model construction. Unlike **Contingent** outcomes, **Probabilistic** outcomes do not depend on system variables, previous tasks, or resource availability, but simply represent a fixed probability of different outcomes. Combining **Contingent** and **Probabilistic** outcomes provides the ability to represent situations where the probability of an outcome depends on system variables, previous tasks, or resource availability. Figure 3 graphically portrays the distinctions between these types of event outcomes.

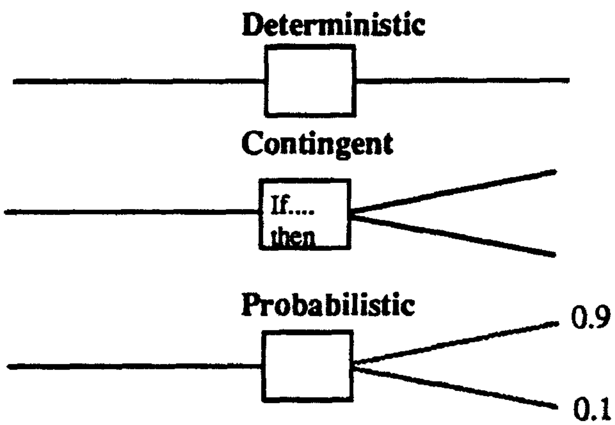


Figure 3. An illustration of the difference between deterministic, contingent, and probabilistic outcomes.

The task of position fixing illustrates the differences between **Deterministic**, **Contingent**, or **Probabilistic** outcomes. For example, a modeling tool that only represents **Deterministic** outcomes would represent a single, invariant, outcome of the task of plotting a course. This outcome might consist of a successfully plotted course, followed by a course change. If a model incorporates **Contingent** outcomes, then the outcome of the course plotting might consist of either a successfully plotted course or an unsuccessfully or inaccurately plotted course, depending on the outcome of previous tasks such as the accuracy of the last position fix. In contrast to the **Contingent** outcome of a task, the outcome of the course plotting task might also be **Probabilistic**. For example, the mate may miscalculate and plot an incorrect course once every 1000 times a course is plotted. In this case, the outcome of the course plotting would be **Probabilistic**, with 999 out of 1000 outcomes being successful course plots and 1 out of 1000 being an unsuccessful or inaccurate plot. Using the example of course plotting, a combined **Contingent** and **Probabilistic** outcome could be used to represent the situation where rough seas and fatigue increase the mate's probability of plotting an incorrect course from 1 out of 1000 to 100 out of 1000.

A.1.7 System-level representation

One major difference in modeling tools lies in their ability to represent the effect of events on system state, and the subsequent effect of these changes on behavior. Three basic distinctions can be made. Describing behavior as a **Task sequence** represents the simplest method of modeling behavior, with no link between the execution of tasks and their effect on the system. The analyst must interpret the consequences of particular sequences on system performance. Other modeling techniques provide the ability to model systems with an **Open loop** representation, or even using a **Closed loop** representation. The **Open loop** representation assumes that behavior proceeds unaffected by changes in system state, while the effect of behavior reflects on system state. Thus, the **Open loop** representation provides a direct measure of how behavior influences system performance. A modeling tool that provides a **Closed loop** representation includes not only the effects of behavior on the system, but also the effect of changes of the system on behavior. Thus, a modeling tool representing behavior as a simple **Task sequence** incorporates no model of the system, while an **Open loop** representation needs only to model the effect of behavior on system performance. A modeling tool that uses a **Closed loop** representation must represent the effects of behavior on the system, as well as a representation of how changes in the system affect behavior.

Figure 4 illustrates the distinctions in system level representations with the task of course plotting. A **Task Sequence** representation would include only information about the mate's ability to generate a series of accurate position fixes and course changes. This information might include the time and accuracy of each task, and gives no information about its effect on system performance. An **Open loop** representation, on the other hand, would include information about ship dynamics, revealing how inaccuracies in position fixing and course plotting affect the ship's actual course. A **Closed loop** representation of the course plotting task would incorporate a model of how inaccuracies in course plotting affect subsequent ship heading, and a model of how these changes affect future course choices. Thus, a **Closed loop** representation would also reveal how the crew adjusts subsequent courses to compensate for earlier errors. Overall, a **Task Sequence** representation provides information on the performance of individual tasks with no explicit implication for system performance, an **Open loop** representation reveals the effect of individual tasks on system performance, and a **Closed loop** representation illustrates not only how individuals influence system performance, but also how system performance influences task execution.

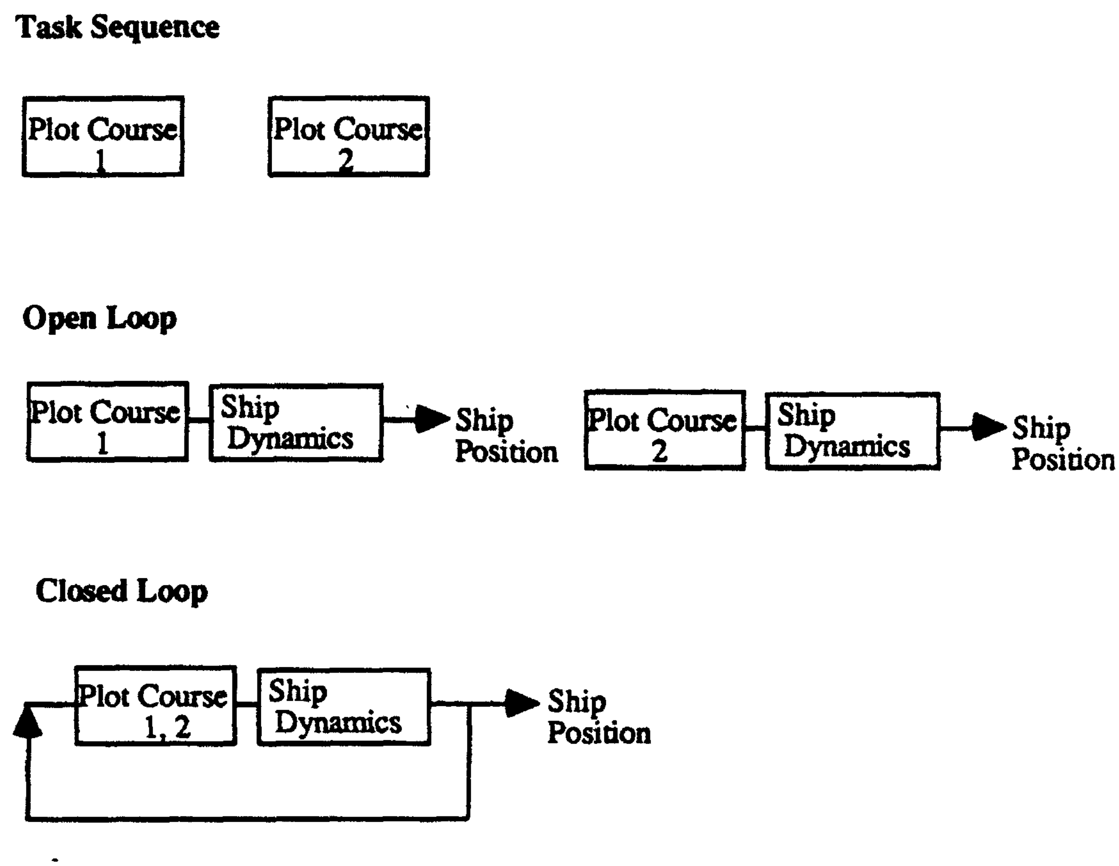


Figure 4. Task sequence, open loop, and closed loop representations of systems.

A.2 COMPARISON OF MODELING TOOLS

Because of the diversity of features available in modeling tools, no one tool contains all possible features. Some tools contain many of the features listed above and others contain relatively few. These differences do not necessarily imply the superiority of any one tool. The value of the tool depends solely on the relationship between the characteristics of the tool and the characteristics of the system under investigation. A tool with many features may be cumbersome to use to investigate a simple system. However, a sim‸ le tool may not contain the features required to represent a complex system accurately. Therefore, the choice of a tool depends on the match between the capabilities of the tool and the characteristics of the system. Some prototypical modeling tools include: spreadsheet, flow diagram, event tree, timeline, task network, control theoretic, queuing theory, and artificial intelligence. The principle features of each of these modeling tools are discussed below.

Spreadsheet: A spreadsheet provides a method of calculating average resource requirements over time. As a manning model, this tool simply multiplies the task duration by the task frequency, sums over all the tasks and divides by the total time available. This calculation provides an estimate of the average time devoted to the task. This technique fails to provide data concerning momentary workload peaks, task sequencing, or performance measures. The output of this model simply relates the task demands (in terms of frequency of occurrence and duration) to the loading of personnel or machines (in terms of the percentage of time occupied by tasks). Although this tool has significant limitations, it can provide useful summaries of time demands of many tasks. In the context of ship manning, this tool could provide one test of sufficient manning (does the average hour total for each crew member meet statutory limits). However, this test ignores workload peaks that might exceed statutory limits.

Flow diagram: A flow diagram charts the progressive travel of materials, information, or people from place to place or from operation to operation (Laughery and Laughery, 1987). This class of tools generally omits a quantitative representation of time, simply representing the sequential ordering of events. In addition, these techniques often do not explicitly identify points in the task sequence that would lead to high workload or increased likelihood of errors. In the context of ship manning, flow diagrams can be used to illustrate the need to coordinate several ship personnel for some tasks. This type of analysis shows where simple summation of workhours of individual ship personnel fails to capture the actual time constraints acting on some personnel.

Event tree: Like a flow diagram, event and fault tree analyses often include the time ordering of events, and do not include the quantitative durations of events. However, an event tree analysis differentiates itself from flow diagrams by describing a system in terms of probabilities of event sequences. An event tree analysis begins at a convenient point in an activity sequence and details the flow of events forward to the fault, while a fault tree begins with a fault and works backward in time revealing the events that led to the fault. The multiple causes and consequences form the branches of the tree. With event/fault tree analysis the emphasis lies on the probabilities that govern the occurrence of different events. Event/fault tree representations specify only event probabilities for a given context, and not time requirements, resource demands, or changes in system performance. This technique could be used in ship manning to calculate system reliability under different manning scenarios.

Timeline: Like flow diagram, a timeline analysis traces the flow of materials or information from place to place or operation to operation. Most frequently a timeline analysis traces the progression of a sequence of tasks, and in some instances resources are coupled to tasks. Unlike a flow diagram, timeline analysis includes an explicit quantitative representation of time. Therefore, this technique includes information about not only task ordering, but also task duration. A timeline analysis also reveals how task demands change over time, illustrating how tasks combine for peak workloads and how task dynamics lead to competition for limited resources.

Task network: Task network techniques combine some characteristics of flow diagrams, event tree analysis, and timeline analysis. In addition, they have added features, such as the ability to incorporate complex contingencies between events, which make the task network approach a powerful modeling tool. Like flow analysis and timeline tools, the task network approach assumes a specific sequence of tasks or events. However, the sequence of events in a task network model may be dependent on a variety of factors, including the outcome of preceding events, system state, or the state of the personnel. While the task network approach offers many features, it represents systems solely as discrete events, failing to consider the effect of continuously changing variables.

Control theoretic: In contrast to the task network approach, the control theoretic approach considers all variables as continuous and does not represent discrete events. In addition, this approach does not model the system as a sequence of tasks, but rather as an interaction between a system to be controlled (the ship) and a controller (the human). In general, the control theoretic approach considers human behavior as an adaptive, goal-directed process, driven by the need to minimize the difference between the current state and a goal state. This perspective has the advantage of explicitly modeling the effect of humans' actions on the performance of the system. The absence of discrete tasks and a representation of resource requirements, combined with the underlying assumption of linearity in the humans perceptual and motor systems, limit the applicability of these models to manual control tasks involving one person.

Queuing theory: Originally developed during World War II to ensure optimum allocation of limited supplies, the field of Operations Research has developed a number of mathematical techniques that support choices between alternatives governed by limited resources. These techniques include task network analysis, simulation, linear and nonlinear programming, and queuing theory. Of these techniques, queuing theory has become a popular technique to represent humans' roles in systems. For example, Rouse and Reid (1981) suggested queuing theory to model the human roles aboard ships. Although queuing theory provides a powerful technique that translates random distributions of tasks over time into measures of system effectiveness and expected levels of workload, the complexities of many real systems make analytic solutions intractable. Possibly the best known and flexible model of queuing theory ($L = \lambda W$) relates the number of waiting tasks (L) to the rate of task occurrence (λ), multiplied by the average waiting time (W). This and other queuing theory models require a number of restrictive assumptions. If the system being modeled fails to conform to these assumptions then the queuing theory models fail to represent the system accurately. Therefore, in many cases queuing theory would fail to accommodate the complexities of shipboard activities, and would not provide output data needed to resolve manning issues.

Artificial intelligence: Artificial intelligence techniques include a wide variety of approaches; however their attempt to provide a framework for modeling human cognition provides the common element linking them. These tools support models that describe how people process information, diagnose problems, resolve ambiguities, and choose between potential actions. These modeling techniques include neural nets, where responses to situations depend on the "activation" of a "node" or combinations of "nodes". An example of this approach includes simulations of human error developed by Reason (1990). More commonly, artificial intelligence techniques simulate human thought by matching the contents of a knowledge base (a representation of operators' knowledge of the system) with the current state of the system. This is accomplished using an inference engine, which represents how humans might process external information. This technique matches environmental situations to preset rules that mimic behavior. Woods, Pople, and Roth (1990) as well as Jones and Mitchell (1991) used this type of technique, employing a blackboard architecture to coordinate knowledge sources and process data to predict operators' intentions. Blackboard architecture is a term that refers to how information from multiple knowledge sources is combined into a global data structure that resolves constraints on behavior implied by the knowledge sources. This architecture is appropriate when simulating a process where multiple, independent sources of information act together to constrain behavior. Overall, artificial intelligence modeling techniques are most important when the aim is to represent the factors that influence cognitive ability. In a manning model this may become important when analyzing the interaction between personnel and automation. If a design assumes that automation heightens decision making ability, eliminating the need for a human in the loop, then an artificial intelligence model could establish the validity of this assumption.

Hybrid techniques: A variety of hybrid techniques exist which combine aspects of several approaches described above. While Table 1 illustrates the capabilities of prototypical examples of these techniques, actual tools may incorporate characteristics of these and other prototypical techniques. For instance, advanced timeline based techniques, such as project management software, incorporate complex contingencies between events and resources. Other tools include so many characteristics common to several tools that they are best labeled as hybrid tools. These tools might combine characteristics of timeline with task network and event tree techniques. By combining the characteristics of several tools, hybrid techniques offer powerful capabilities for developing manning models. MicroSAINT (Laughery, Drews, and Archer, 1986) provides an example of a technique that most closely resembles a task network approach; however, it includes many of the capabilities of spreadsheet, flow diagrams, and event tree approaches. Enhancing the task network technique by incorporating such capabilities as complex probabilistic branching and continuous variables, provides MicroSAINT with greater capability to represent complex systems more accurately.

Like MicroSAINT, the modeling tool adopted for PROCRU (Baron et al., 1980; see section 2.1) represents a hybrid tool, combining aspects of an artificial intelligence approach with a control theoretic approach. PROCRU expands on the techniques of modern control theory to include some of the characteristics of artificial intelligence modeling tools. Including these functions facilitates creation of models that not only capture the ability of personnel to perform continuous control tasks, but also the ability to select and execute procedures that include a series of discrete activities.

A.3 CONCLUSIONS

Modeling tools partially determine the usefulness of a manning model by limiting the types of and nature of relationships that a model can include. Table 1 summarizes the limits and capabilities of a variety of modeling tools. If the model capabilities fail to match the characteristics of the system being modeled then model validity will suffer. Therefore, models that include more filled dots offer greater capability for modeling a wide variety of systems. While the task network approach supports a wide variety of modeling capabilities, the diversity of capabilities alone does not guarantee success. For instance, the timeline approach has far fewer capabilities, but it includes the ability to model continuous variables, which could be essential in modeling some systems. This comparison shows that the number of dots on Table 1 alone is not a good measure of the value of a modeling tool. Instead, the worth of a tool depends on how closely tool characteristics match system characteristics. Thus, Table 1 could be used as a checklist, verifying that the tool can model the system.

Further, there will be differences in the expertise and time required to develop a model, depending on the modeling tool. Some tools offer simple intuitive interfaces that enable novices to enter data and create a model. Other tools require much specialized knowledge. The spreadsheet, flow diagrams, event trees and timelines require little training for model development. These tools require only a thorough knowledge of the specific system, modeling formalisms, and limits of the tools. Task network and control theoretic modeling approaches require substantial knowledge of the specific tool. This specific knowledge compares to learning a programming language. Models relying on control theoretic, queuing theory and artificial intelligence all require substantial expertise related to the theoretical background underlying these approaches. In addition, the complexity of these approaches requires much more development time, compared to other approaches. Thus, the order of modeling tools in Table 1 provides a rough metric of the expertise and time required to create a model with each of these approaches. In other words, by presenting the models in a left to right sequence, starting with the spreadsheet and ending with the artificial intelligence approach, Table 1 indicates knowledge required to use each of these tools.

Table 1. The relationship between modeling tool capabilities of modeling tools and tool characteristics.

Tool Types / Tool Characteristics	Spread-sheet	Flow Diagrams	Event Trees	Timeline	Task Network	Control Theoretic	Queuing Theory	Artificial Intel-gence
VARIABLE REPRESENTATION								
Continuous				●		●		○
Discrete	●	●	●	●	●		●	●
Combined				●				○
TASK ORDERING								
Concurrency		●		●	●		●	
Priority					●		●	●
Deterministic	●	●	●	●	●		●	●
Contingent		●	●		●	●	●	●
Probabilistic			●		●	●	●	○
EVENT TIMING								
Deterministic	●		●	●	●		●	○
Contingent			●		●	●		○
Probabilistic			○		●	●	●	○
RESOURCE REQUIREMENTS								
Deterministic	●	●		●	●	○	●	
Contingent					●	○		
Probabilistic					●	○	●	
RELATION TO SYSTEM								
Task sequence	●	●	●	●	●		●	●
Open loop			○		●	●	●	
Closed loop					○	●		

● Tool accommodates indicated dimension.

○ Tool marginally accommodates indicated dimension.

APPENDIX B FEATURES OF MANNING MODEL CONTENT

This Appendix discusses how the type of information incorporated into a model guides data collection and determines the type of information the model will produce. Similar to the great variety of modeling techniques, a model of the ship system may incorporate many different types of information. The content of a model may range from a simple description, limited to a small part of a ship, to a very detailed description that encompasses an entire ship in great detail. As with the different modeling techniques, different types of information incorporated into the model will influence both the accuracy and cost of the model. Therefore, the capabilities and costs of a manning model depend not only on the modeling tool, but also on the model content.

In general, a model becomes more accurate as the content of the model becomes more inclusive, matching the actual system more closely. However, the more information the model contains, the more complex and costly it becomes. Since no model perfectly replicates the actual system, the content of a manning model should depend on the specific questions that the model should address. This chapter provides a classification, based on six dimensions, of the different types of information that a model might incorporate. Each of these six dimensions can describe each of the two primary components of a manning model. The first of these components is a description of **task demands** (what personnel must do) imposed on ship personnel. These task demands depend on the particular trade route, ship configuration, maintenance program, level of automation, or emergency situation. The second component of a manning model is the **profile of factors** that relate the task demands to the safety, efficiency, and legal constraints associated with ship operation. These parameters might include human performance limitations such as fatigue, memory and strength, as well as the effects of high mental workload and the limits of decision making effectiveness. Besides human performance limits, a manning model might also include factors that relate task demands to the rules, regulations, and guidelines that govern the maritime community. For example, these factors might show that task demands require the crew to violate the 12 hour work day rule imposed by the Oil Pollution Act of 1990. A single model might include both human performance and regulatory limitations; however, constructing two models, one based on human performance and the other based solely on the legal constraints, could provide a clear illustration of regulatory deficiencies. While all the dimensions apply to description of both the task demands and the profile of factors, some dimensions are more important in the description of task demands, and others are more important for the description of the profile of factors. In summary, each dimension can describe the task demands imposed on ship personnel, as well as the profile of factors that overlay these demands to reveal the legal, safety, and efficiency consequences for the system.

B.1 DESCRIPTION OF CONTENT DIMENSIONS

B.1.1 Scope of analysis (Inclusiveness)

The scope of the model refers to how inclusive a model is, specifying what aspects of the complete system (crew, ship, or shore support) the model addresses. In other words, the scope of the model specifies the "system" under examination. In some instances the "system" may be defined very narrowly and in others it may be defined quite broadly. For example, the scope of a model specifies the aspect or aspects of the ship included in a model (bridge, engine room, catering facilities), the phase of the voyage (port operations, open ocean, piloted waters), or the type of operation (fire fighting, spill mitigation, off loading, collision avoidance). Narrowing the scope simplifies a model, but it may also reduce model validity and applicability. For example, a manning model developed with a scope narrowed to open ocean operations may provide misleading estimates of crew size for fire fighting and port operations. At the same time, a model whose scope covers not only the entire ship, but also all phases of the voyage, under all types of operations, would require enormous resources to construct. In general, limiting the scope of the model decreases validity, while increasing the scope of the model increases the resources required to construct the model. While a wide variety of dimensions might define the scope of model content, some of the most salient include: aspect of ship, type of ship, phase of voyage, subset of crew, and type of operation. Table 1 provides examples of the content of possible models and various dimensions that define the scope of the analysis of manning models.

B.1.2 Level of detail/time scale (Specificity)

Like the scope of analysis, the level of detail of information included in a model acts to determine model validity and cost. A greater level of detail may reflect the behavior of the system more accurately, but model development may involve more effort and expense. The level of detail defines the amount of information the model contains about either events across time or aspects of the system. Information incorporated into a model might include extremely detailed information or a simpler aggregation of detailed information into a coarser description of the system. Aggregating information might summarize task elements over time or it might amalgamate different aspects of the physical system. For example, a model might include information about position fixing, course plotting, and course adjustment, or it might aggregate these individual tasks into a summary task such as "navigation." Likewise, for the physical system, the level of detail may shift from including information about individual pipes and valves involved in an oil tanker's unloading, to an aggregate that represents all the individual components as a single piping system. Thus, less detail results in a simpler model, but it also limits how closely the model can mimic the actual system.

Table 1. **Dimensions that limit the scope of a manning model and examples of model content for each of these dimensions.**

Aspect of ship	Aspect of voyage	Subset of personnel	Type of operation
Bridge	Open ocean sailing	Captain and pilot	Fire fighting
Engine room	Piloted waters	Bridge staff	Man overboard
Deck	Heavy weather sailing	Day workers	Bunkering
Control room	Docking		Oil spill mitigation
			Cargo loading
			Navigation

B.1.3 Interactions among subsystems

An important consideration in defining the scope of the model involves how the model incorporates interactions across parts of the system. This issue becomes important because many system components interact with other components. However, most of these interactions are extremely small, and model accuracy does not suffer by their elimination. Since model scope defines the parts of the system included, it eliminates all interactions with parts of the system not included in the model. If significant interactions exist, but are not incorporated into the model, then the model will not accurately reflect the system and it may produce misleading results. The development of a model must strike a balance between the extreme systems perspective, in which the model must include the entire system because all aspects of the system depend on all other aspects, and a reductionistic view that sees all aspects of the system as independent.

The ability of a mate to plot a course and communicate with the VTS provides a concrete example illustrating the potential interactions across aspects of the ship/shore support system. A model of the mate's performance managing communications might operate independently of another model of his performance plotting the course. Alternatively, a single model might be constructed that predicts both performance on the course plotting and the communication with the VTS. A comparison of these two modeling approaches reveals the potential problems with each. The results of the first would be very misleading if calls from the VTS interrupt the mate's task of course plotting, leading to errors in both the course plotting and VTS communications. The complexity of the second model may enable it to predict performance of VTS communications and course plotting accurately, but it may require more resources to develop and it may be more difficult to use.

B.1.4 Level of abstraction

Unlike the level of detail and scope, changes in the level of abstraction address the focus of the model rather than the tradeoff between validity and costs. The level of abstraction specifies the types of terms and concepts used to describe the system. These terms and concepts determine what data the model will require, as well as what information the model will provide. The level of abstraction varies from concrete descriptions that include information about the physical characteristics of the system (distances, colors, sizes), to more abstract representations that describe the system in terms more closely related to the functional characteristics or purposes of the system (information flow, transfer functions, and functional relationships). Figure 1 provides a graphical illustration of the same system at two levels of abstraction. The physical level includes information concerning the physical configuration of the ship, while the functional level includes information about the functional linkages between parts of the ship.

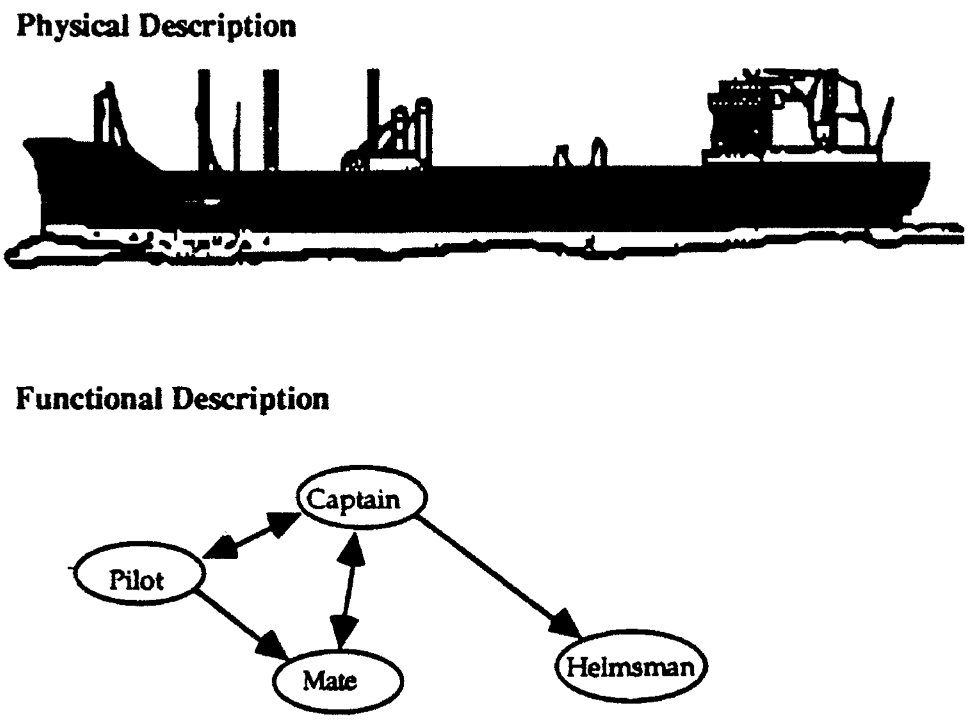


Figure 1. An illustration showing that a low level of abstraction (Physical), which includes the physical characteristics of the ship (distances and physical positions) provides different information compared to a more abstract description (Functional), based on information flows and authority gradients.

Models developed using different levels of abstraction provide answers to different kinds of questions. For example, a model designed to estimate the effectiveness of the crew to monitor the filling of cargo tanks requires a model that includes information at a relatively low level of abstraction. Since the ability to monitor the tanks depends on their ability to move from tank to tank, the physical location of the tanks will limit the frequency that the personnel would be able to check the tanks. If the model did not include information at a low level of abstraction, describing the physical location of the tanks, then it would fail to provide information about how different loading procedures affect the ability to monitor the level of the tanks and prevent spills due to overfilling.

Some questions require more abstract information, and the data about the physical configuration of the ship only complicates the model. For example, in modeling communications between pilot, captain, tugs, dock workers, and deck personnel during the docking of a ship, the more abstract concepts of information flows and authority gradients provide a useful description of a system.

B.1.5 Veridicality of content

Veridicality of model content refers to the accuracy with which the model represents the cognitive and physical limitations of humans. At one extreme, the model may consider human limits only in terms of workhour requirements. At the other extreme, the model may include information about not only the decision process, but also the physical limits of perception and biomechanics. In a manning model, both physical and cognitive limits of humans may be important. For example, physical demands as well as mental demands influence the fatigue of personnel, and fatigue may lead to an increase in human error, thus reducing ship safety. A manning model may contain a wide variety of cognitive and physical characteristics that limit the ability of personnel to control the system. Some of these factors include limits in auditory and visual capabilities, as well as memory capacity and crew coordination. Including the cognitive and physical limits of humans acts to increase the veridicality of the model; however, they will also increase its cost and complexity.

In addition to including the cognitive and physical characteristics that influence human performance, model veridicality also depends on the representation of the decision making process. Just as a more veridical model may result by including human cognitive and physical characteristics, the representation of the decision making process can vary from simply a restatement of the task demands, to a model based on a psychological theory of how humans integrate information. At one extreme, the model includes only information concerning what personnel must do to satisfy the demands of the system, and at the other extreme, the model includes detailed information about cognitive processes involved in decision making.

Three major approaches differentiate between methods used to represent the decision making process: **Normative, Outcome,** and **Process. Normative** modeling represents decision making in terms of a normative procedure. That is, the model would represent decision making as if decisions conform to a mathematical ideal, optimally satisfying the demands of the task. Edwards (1962) introduced this approach by using a Bayesian approach to combine statistical information about possible decision choices to predict the optimal choice where optimal decisions depend on a global estimate of the probabilities of failure or expected gain of the possible choices. Basing a model of behavior on emergency procedures provides another example of the normative approach. In this case emergency procedures are normative in the sense that they represent the proper coarse of action. The second approach, outcome modeling, represents decision making by simply matching the outcome of the model with that of actual operators. This approach addresses only the correlation between the behavior of the system being modeled and the output of the model. The ability of linear regression equations to predict human decision making performance illustrates the utility of this approach (Dawes, Faust, and Meehl, 1989). The third approach, **process** modeling, represents the actual decision making process, incorporating a description of the steps involved in the decision making (Payne, 1976; Payne, Bettman, and Johnson, 1988). This description might include strategies, and the cognitive limitations and capabilities that constrain the decision making process. This modeling approach includes a description of the cognitive processes involved in the decision making task, and provides an outcome measure of system performance, as well as an understanding of how the process might fail.

It is also important to consider the overall relationship between the crew and the system. Increasingly sophisticated technology has changed the role of the human, distancing personnel from the process through increasing levels of automation. As more sophisticated technology is implemented personnel act more as supervisory controllers, rather than direct manual controllers (Sheridan, 1987). In this situation, personnel spend little time actively involved in the process, but they monitor the automation and intervene when the automation fails. This presents a particularly difficult situation to model. Workload may vary greatly, dropping when automation carries out the task and the human simply monitors its performance, and increasing when the human intervenes and returns to manual control. A model that fails to incorporate the factors that influence transitions between automatic and manual control may produce estimates of workload and crew requirements that differ greatly from reality (Sheridan and Hennessy, 1984). Simply assuming that ship personnel will take advantage of the capabilities of the automation may underestimate the crew required to staff the ship adequately. Thus, veridicality of a model's representation of personnel may vary from including a detailed description of how the demands of supervisory control influence decision making, to a model that ignores all the physical and psychological characteristics of personnel.

B.1.6 Organization of model content

The organization of model content defines the types of questions that the model might answer. Sections B.1.1 to B.1.5 have described how models might differ in their content, and this section describes the organization of model content. The variables associated with a model fall into three categories: **Parametric/Input, Structural,** and **Outcome.** The **Parametric/Input** variables represent aspects of the system that are manipulated to determine their effect on system performance. In the context of a manning model, **Parametric/Input** variables might include crew size, trade route, weather, traffic density, crew expertise, and organizational structure. Adjusting these variables reveals how these factors affect the operation of the system.

In contrast to **Parametric/Input** variables, **Structural** variables represent the constraints and demands of the system. These variables are fixed and represent relationships contained in a manning model. These variables might include: the effect of fatigue on the time to perform a task, the memory limitations of crew members, the tasks involved, the relationship between behavior and system state, and the probability of successfully completing tasks. Structural variables comprise the model content, and would remain unchanged during model use.

Outcome variables represent measures of system performance. These variables include the overall reliability of the system, the number of neglected tasks, the number of hours on duty, or the precision of a maneuver. The complexity of the model content and the choice of modeling tool limit the outcome variables of the model. For example, a timeline model cannot produce estimates of overall system reliability. Instead, a timeline-based approach would illustrate which tasks might not be performed because of multiple, simultaneous demands on a single person's time. Likewise, an event tree diagram, which represents the system with event probabilities, might fail to capture the effects of high levels of workload, but it would provide an estimate of the overall system reliability. Because the choice of model content and modeling tool defines the possible outcome measures, selecting the most appropriate tool and content is critical.

B.2 CONCLUSIONS

Just as characteristics of tools limit the capabilities of manning models, model content limits the ability of a model to mimic reality. The detailed description of six dimensions of model content shows how each of these dimensions affects model accuracy. Some of these dimensions (scope of analysis and level of detail) refer to the breadth of a model's description of a system. Other dimensions, such as the interactions among subsystems, level of abstraction, and veridicality of content, refer to the types of relationships included in a model. The final dimension, organization of model content, illustrates the limits of the model determined by how information is incorporated into the model. This organization specifies which variables may be manipulated and which are fixed. These six dimensions of model content combine with the limits and capabilities of a modeling tool to determine model validity and the type of information that a model can produce. Therefore, creating a manning model should involve an evaluation of how model specifications affect each dimension, and how these effects influence model utility.