NUREG/CR-2669
ORNL/TM-8300

ornl

OAK RIDGE NATIONAL LABORATORY

UNION CARBIDE

Front-End Analysis for the Nuclear Power Plant Maintenance Personnel Reliability Model

Arthur I. Siegel
William D. Bartter
J. Jay Wolf
Helmut E. Knee
Paul M. Haas

This Work Performed For
Nuclear Regulatory Commission
Under
DOE Interagency Agreement 40-550-75

OPERATED BY
UNION CARBIDE CORPORATION
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161

Available from
GPO Sales Program
Division of Technical Information and Document Control
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

NUREG/CR-2669
ORNL/TM-8300
Distribution Category RX

Contract No. W-7405-eng-26
Engineering Physics Division

FRONT-END ANALYSIS FOR THE NUCLEAR POWER PLANT MAINTENANCE PERSONNEL RELIABILITY MODEL

Arthur I. Siegel*
William D. Bartter*
J. Jay Wolf*
Helmut E. Knee
Paul M. Haas

NRC Monitor: Thomas G. Ryan
Human Factors Branch
Division of Facility Operations

Program Manager: P. M. Haas
Oak Ridge National Laboratory

Principal Investigator: H. E. Knee
Oak Ridge National Laboratory

Manuscript Completed: March 1983
Date Published: August 1983

This Work Performed For
Nuclear Regulatory Commission
Under
DOE Interagency Agreement 40-550-75
NRC FIN No. B0461

Work Performed Under
ORNL Subcontract No. 40X-70527

*Applied Psychological Services, Inc.
Science Center
Wayne, Pennsylvania 19087

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37830
operated by
UNION CARBIDE CORPORATION
for the
DEPARTMENT OF ENERGY

TABLE OF CONTENTS

Page

FOREWORD

This report is one of a series that is planned to describe the results of a program undertaken by the Oak Ridge National Laboratory (ORNL) for the U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, to define, develop, validate, and disseminate a methodology for the quantitative prediction of human reliability in the conduct of maintenance tasks in nuclear power plants (NPPs). ORNL has subcontracted portions of this effort to Applied Psychological Services, Inc. A program scoping/feasibility study has been completed which consisted of four tasks:

1. Completion of a structured user survey to assess the likely utility of the proposed methodology, to identify user output requirements, and to identify critical input data needs.
2. Assessment of available methodologies via completion of a literature survey.
3. Completion of the initial job analyses of key maintenance positions.
4. Definition of a comprehensive program plan for development, validation, and dissemination of the proposed methodology.

Three NUREG/CR reports are planned from this scoping/feasibility study: (1) this report (NUREG/CR-2669), a summary report on the program planning effort, describing the front-end user requirements analysis, literature review, and program plan, (2) a report on the job analysis of the maintenance mechanic position (NUREG/CR-2670), and (3) a report on the job analysis of the instrument and control supervisor and maintenance supervisor positions (NUREG/CR-2668).

The job analyses of the instrument and control technician position and the electrical position have been completed and will be summarized in subsequent NUREG/CR reports (NUREG/CR-3274, ORNL/TM-8754 and NUREG/CR-3275, ORNL/TM-8755, respectively). In addition, reports will also be prepared which will address the model development, sensitivity testing and validation efforts.

H. E. Knee, ORNL, Principal Investigator

EXECUTIVE SUMMARY

Problem

Human errors committed during maintenance activities have been potentially major contributors to the overall risk associated with the operation of a nuclear power plant (NPP). In order to ascertain the impact of maintenance personnel performance on overall plant safety/reliability/availability and to assess the effects of proposed maintenance improvements prior to their actual implementation within a NPP, it is desirable to have a structured, predictive methodology which addresses NPP maintenance personnel activities that has been validated for use with NPP systems. Such a methodology would provide information to aid in establishing standards, assessing risk/safety, and increasing plant maintainability.

Tasks

The front-end user survey consisted of a structured interview which inquired into factors influencing nuclear power plant maintenance. It was administered to 30 individuals from nuclear power plants, the Nuclear Regulatory Commission, and architect/engineer organizations. The results of these interviews provided an initial indication of the usefulness of various types of information which could be provided by a structured methodology. The interview results also served as a basis for a questionnaire which was completed by individuals from 68 nuclear power plants, Nuclear Regulatory Commission, and architect/engineer representatives.

The literature review focused on the two major classes of existing human reliability techniques — analytic techniques and simulation techniques. From information obtained from the front-end user survey and from a compilation of desirable characteristics listed later in this report, simulation techniques were chosen as the type of methodology to be developed within this program.

Once simulation modeling had been chosen, a detailed program plan was developed which outlined three subsequent phases for this program. The development phase, which spans a 20-month period contains 4 tasks, the validation phase, which spans 16 months, contains 5 tasks, and the dissemination phase which spans 10 months contains 2 tasks. The total program plan will be accomplished in 38 months and will result in a fully validated simulation model that will be applicable to NPP maintenance activities.

Results of the Front-End User Survey

The structured interviews and mail survey results indicated that the potential users of a structured methodology for maintenance would find much of the information that could be provided by such a technique "extremely useful" or "very useful." The types of information with the highest usefulness ratings were:

- the relationship between training and performance proficiency
- information about the presence and causes of undetected errors
- information about maintenance team composition

- information about time to complete maintenance tasks and about error rates
- information about methods for minimizing radiation exposure levels

The questionnaire also surveyed opinions about causes and consequences of maintenance error. The majority of the respondents believed that:

- fault diagnosis is the maintenance task most likely to be incorrectly performed
- the presence of time constraints is the greatest error contributing factor
- errors are usually manifested because technicians do not follow procedures
- errors made by instrument and control technicians are the most likely to involve an element of public risk

Conclusions of the Front-End User Survey

The implications of the results were:

- a computer simulation model would prove very useful to a variety of public protective and interest purposes
- such a model should provide a wide variety of output data and be rich in internal variables
- the internal variables should include, but not be limited to, environmental conditions, maintainer ability, physical factors, and performance moderating (e.g., stress, fatigue, heat) variables
- the model should be applicable to the maintenance activities of all nuclear power plants and should yield output at both the subtask and the task levels

Recommendations

Because of (1) the potential of such a model for nuclear power plant maintenance analytic purposes, (2) the perceived utility of such a model as expressed by the participants in the survey and by the interviewees, and (3) the relatively low risk of a model development effort, completion of initial work towards the design of a nuclear power plant maintainer performance reliability model is recommended. From the needs identified within the front-end user survey and the results of the literature review, the recommended model should be of the simulation type. This simulation model should be sensitivity tested, validated, and eventually disseminated to identified potential users. A plan for the accomplishment of these tasks is presented in this report.

LIST OF FIGURES

LIST OF TABLES

Table Page

ABSTRACT

The front-end analysis performed for the Nuclear Power Plant Maintenance Personnel Reliability Modeling Program consisted of three primary tasks which are addressed within this report. The first of these was a front-end user survey which investigated the need for and potential content of a structured methodology for nuclear power plant maintenance. For this survey, personal interviews and mail surveys were completed to collect the needed information. A large percentage of the interviewees (80%) said that a structured methodology would be "extremely useful" or "very useful." The classes of information rated most highly by potential model users were: training, undetected errors, radiation exposure, time required for task completion, and proficiency.

The second task of the front-end analysis assessed available human behavioral methodologies for applicability and adaptability potential for nuclear power plant maintenance activities. Although a large number of human behavioral methodologies were found to exist, many were very narrow in scope, had little applicability to maintenance tasks, and did not supply the types of information identified through the front-end user survey. Evaluation of the existing human behavioral methodologies coupled with the needs identified through the front-end user survey led to the choice of simulation modeling as the type of methodology to be developed.

The third task of the front-end analysis was the development of a comprehensive program plan for the simulation model. The plan that was developed spans a 38-month period, and includes three phases: the model development phase, which will result in the release of a debugged/sensitivity tested/calibrated version of the model; the validation phase, which will result in the release of a validated version of the model; and the dissemination phase.

1. FRONT-END USER SURVEY

1.1. Semistructured and Structured Interview Development and Administration

As an introductory step, a set of preliminary interviews was conducted with nuclear power plant personnel. The results of these interviews provided the basis for a more extensive set of structured interviews with Nuclear Regulatory Commission, nuclear power plant, and architectural and engineering organization personnel. The results of this work, in turn, provided the structural and content basis for a mail survey with a larger sample of persons for whom a structured methodology may possess utility. Section 1 of this report presents the methods of interview development and a discussion of the interview findings.

1.1.1. Introductory Semistructured Interviews

Initially, a set of areas for discussion and investigation during the semistructured interviews was developed on the basis of analysis of relevant documents and as the result of conversations with knowledgeable persons. Examples of the documents reviewed are: Finnegan *et al.*,[1] Hall *et al.*,[2] IEEE Task Group,[3] Luckas and Hall,[4] and Swain and Guttman.[5]

Appendix A presents the topic area list which provided the basis for focusing the initial interviews.

1.1.2. Interview Administration

The introductory interviews were administered at two nuclear power plants. The content was then revised and the administration was fully structured. The interview was administered in this structured form (Appendix B) to 30 interviewees including Nuclear Regulatory Commission (NRC), nuclear power plant, and architectural and engineering organization personnel. Table 1.1 presents the number of interviewees in each of these categories.

The interviewers were four experienced human factors specialists with depth of experience in interview methods and procedures. In its final form, administrative time was 1.5 hours on the average.

Each interview was conducted individually and privately. Interviewees were encouraged to elaborate on their answers so that full insight could be obtained.

1.2. Interview Results

One purpose of the interviews was to provide a test-bed for the interview questions which would subsequently be incorporated into a questionnaire for wider distribution. The 30

Table 1.1. Number of Interviewees in Each Type of Organization

Nuclear Power Plant Type	Number of Interviewees
PWR-1	2
BWR-1	4
PWR-2	4
PWR-3	4
PWR-4	4
Architectural and Engineering Companies	
Company 1, NY	4
Company 2, PA	4
Regulatory Agencies	
Nuclear Regulatory Commission, Washington, DC (NRC)	4
TOTAL	30

interviews that were conducted are considered to be sufficient for that purpose. However, with 30 interviews, the findings can be considered as indicative but not definitive.

The responses were collapsed across the three types of interviewees so that confidence in the results could be as great as possible. The results justify this collapsing because there were no response differences among the three groups for the majority of questions. However, those items on which differences were evidenced will be discussed in some detail.

1.2.1. Results

Table 1.2 summarizes the responses to a majority of the items included in the interview. These items dealt with the usefulness of various types of information that could be provided by a structured methodology. The results are given in terms of the percentage of interviewees who indicated that the information could be (1) "extremely useful" or "very useful"; (2) "useful"; or (3) "slightly useful" or "not useful." Although the interviewees could rate the usefulness of information on a five category scale, the extreme ratings were collapsed to present the data more clearly. The present discussion concentrates on the five items which had the highest "useful" rating (most useful information) and the four items which had the highest "not useful" rating (least useful information).

Table 1.2. **Summary of Responses to Selected Items in Interview**

Item	Topic	Percentage Responding: "Extremely Useful" and "Very Useful"	"Useful"	"Slightly Useful" and "Not Useful"
11	Task time allowance	53	27	20
12	Team composition	57	37	7
13	Time adjustments	47	33	20
14	Team proficiency	47	43	10
15	Outside contractors	41	28	31
16	Undetected error	70	10	20
17	Noise effects	47	33	20
18	Temperature effects	50	33	17
19	Safety gear	57	20	23
20	Tight places	50	33	17
21	Time constraints	57	30	13
32	Stress	53	33	13
33	Morale	52	38	10
34	Training	80	20	0
35	Proficiency	65	19	16
55	Subtask success probability	60	23	17
56	Time of day	53	23	23
57	Fatigue	60	30	10
63	Risk	57	20	23
66	Demand vs. preventive maintenance	40	50	10
71	Safety	67	17	17
72	Human reliability	53	30	17
74	Overall usefulness	80	20	0

1.2.2. Most Useful Information

The item which received the highest overall rating was item 74, which asked for a general rating of the usefulness of a computer oriented technique. Eighty percent of the interviewees said that such a model would be either "extremely useful" or "very useful." There were no differences in the three types of respondents (plant personnel, NRC, and architectural and engineering personnel).

Item 34 was equally favorably rated. This item asked about the usefulness of information on the relationship between training and the likelihood of successful task completion. Again, 80 percent of the respondents rated information of this type as either "extremely useful" or "very useful." Again, there were no group differences.

Information on the likelihood of undetected errors (item 16) was rated as "extremely useful" or "very useful" by 70 percent of the interviewees. There were no differences between the three groups on this item.

Information about safety hazards to maintenance personnel (item 71) was also highly rated. Of the 30 interviewees, 67 percent rated this type of information as being either "extremely useful" or "very useful." Although the numbers are small, it is noteworthy that of the 18 power plant respondents, 5 (28 percent) rated information on safety hazards as being only "slightly useful" or "not useful" while none of the 12 respondents from architectural and engineering organizations and the NRC gave this information such a low rating.

Finally, information relating the maintenance technician's proficiency and the probability of successful task completion was rated as "extremely useful" or "very useful" by 65 percent of the respondents. There were no differences in response pattern among the three types of respondents.

1.2.3. Least Useful Information

Table 1.2 also shows the percentage of the interviewees who thought the various items of potential information that could be provided by a model would be only "slightly useful" or "not useful." In general, these values are low and range from zero to 31 percent. The four items with the largest "slightly useful" or "not useful" ratings are briefly reviewed.

Information on trade-offs between contracting work out and performing it by plant personnel (item 15) was rated as "slightly useful" or "not useful" by 31 percent of the interviewees. It was not surprising that fewer representatives of power plants found this information "slightly useful" or "not useful" (18 percent) than architectural and engineering organization representatives (38 percent) or NRC representatives (75 percent). It is clear that this information became less interesting to an interviewee as his distance from a power plant increased.

Information about the effects of wearing safety gear on performance (item 19), information on time of day effects (item 6), and information on risk (item 63) was rated "slightly useful" or "not useful" by 23 percent of the respondents. There were no group differences on the safety gear and time of day issues among the three groups. But, on the risk item, there were marked differences. Of the nuclear power plant respondents, 17 percent rated this information "slightly useful" or "not useful." The corresponding values for the architectural and engineering organizations and the NRC interviewees were respectively 51 percent and zero percent. In fact, 100 percent of the regulatory agency interviewees rated information on risk as "extremely useful" or "very useful."

1.2.4. Usefulness Ranking

Two items were included in the interview which probed into the usefulness of information about: (1) the effects of selected physical factors on performance, and (2) the effects of selected "psychological" factors on performance. In both items, the interviewees were asked to rank order various factors in terms of the usefulness, to them, of information about the factors. The results from these two items are presented in Tables 1.3 and 1.4. Heat effects and training effects were respectively ranked first, while time constraint and stress effects were respectively ranked second.

Table 1.3. Ranking of Usefulness of Information on Physical Factors

Factor	Rank 1	2	3	4	5	Mean Rank	Final Order
Noise	0	13	10	23	53	4.13	(5)
Heat	43	13	20	23	0	2.21	(1)
Safety gear	13	20	30	23	13	3.00	(4)
Time constrants	27	17	27	13	17	2.79	(2)
Tight places	17	37	13	17	17	2.83	(3)

Note: Entries in cells represent percentage of responses.

Table 1.4. Ranking of Usefulness of Psychological Factors

Factor	Rank 1	2	3	4	Mean Rank	Final Order
Stress	23	23	33	20	2.48	(2)
Training	57	13	27	0	1.76	(1)
Morale	10	27	20	43	2.96	(4)
Proficiency	10	37	20	33	2.76	(3)

Note: Entries in cells represent percentage of responses.

1.2.5. Model Content

Table 1.5 summarizes the results from nine interview items that related to the content and structure of a computerized model for predicting maintenance technician performance. As with the prior analyses, the results are collapsed across the three groups of respondents. From the point of view of model content, the results suggest the inclusion, as a minimum, within any structured methodology, the ability to address fault diagnosis, preventive maintenance, radiation exposure levels, a number of personnel ability traits, and subtask as well as task performance. In most instances, the pattern of responses to the content oriented questions was the same for the three groups.

In one question, the interviewees were asked to indicate the types of maintenance tasks most likely to be performed incorrectly. Overall, the two most frequently mentioned tasks were fault diagnosis (40 percent) and preventive maintenance (20 percent). However, when the response patterns of the three groups were examined separately, a number of

Table 1.5. Summary of Responses to Model Content Oriented Questions

Topic	Percentage Responding
Task most frequently performed incorrectly	
Fault diagnosis	40
Repair	16
Replacement	3
Checkout	7
Calibration	2
Preventive maintenance	20
Inspection	10
Other	2
Importance of information on radiation exposure level	
Important	93
Unimportant	3
No Response	3
Usefulness of radiation exposure information	
Extremely Useful and Very Useful	69
Useful	31
Slightly Useful and Not Useful	0
Factors important in selection	
Intelligence	26
Mechanical ability	19
Determination to get things done	11
Ambition	2
Prior experience	18
Precision (accuracy) in performing a task	18
Other	6
Factors important on the job	
Intelligence	16
Mechanical ability	11
Determination to get things done	21
Ambition	3
Prior experience	11
Precision (accuracy) in performing a task	31
Other	7
Information on errors	
Given by task	20
Given by subtask	30
Both	50
Form of model output	
Printout	74
Cathode ray tube	26

differences appeared. For respondents from nuclear power plants, fault diagnosis was the single most frequently mentioned item (47 percent). Repair was mentioned second most frequently (18 percent), followed by inspection (15 percent). For these respondents, preventive maintenance was the item least likely to be mentioned. Respondents from architectural and engineering organizations, on the other hand, indicated that preventive maintenance is the area most likely to produce errors. All eight of these respondents mentioned preventive maintenance, whereas only one of the 18 power plant respondents so replied. Finally, two of the four respondents from the NRC responded "preventive maintenance" to this item. This pattern suggests clear differences of opinion in "where mistakes are made in nuclear power plant maintenance." As mentioned before, due to the comparatively small sample size, these findings should be considered to be only suggestive.

Item 65 also evoked replies which suggested some differences among the three groups. Item 65, like item 8, was concerned with preventive maintenance and asked for an estimate of the percentage of preventive maintenance tasks that are associated with public risk. All of the respondents from the NRC said that the percentage is less than 50 percent. The responses of the power plant and architectural and engineering organization representatives contrasted with this consistent pattern. Of these respondents, 65 percent indicated that risk is associated with less than 50 percent of the preventive maintenance tasks, and 35 percent of these respondents said that risk was associated with over 50 percent of the preventive maintenance tasks. As for all of the responses to this interview, these differences are differences in opinion, not differences in fact. However, these opinions may be related to operating and policy decisions.

Tables 1.6, 1.7, and 1.8 summarize the responses to three items in which respondents were asked to rank various factors vis-a-vis their contribution to errors in general, undetected errors, and the perceived contribution of various personnel types to public risk. "Heat" and "not following procedures" were ranked highest as contributing to one or the other type of error, and the respondents perceived that errors made by instrument and control technicians (Table 1.8) were the most likely to lead to an increase to public risk. Mechanical technicians and electrical technicians followed instrument and control technicians in that order.

Table 1.6. Ranking of Factors by Contribution to Errors in General

	Rank						
Factor	1	2	3	4	5	Mean Rank	Final Order
Noise	0	13	17	20	50	4.07	(5)
Heat	50	17	23	10	0	1.93	(1)
Safety gear	10	30	23	17	20	3.07	(3)
Time constrants	17	10	23	30	20	3.26	(4)
Tight places	23	30	13	23	10	2.64	(2)

Note: Entries in cells represent percentage of responses.

Table 1.7. **Ranking of Factors Contributing to Undetected Errors**

Factor	Rank 1	2	3	4	Mean Rank	Final Order
Poor system design	24	10	41	24	2.63	(3)
Inappropriate tools	0	7	32	61	3.54	(4)
Not following procedures	57	17	23	3	1.72	(1)
Poor procedures	14	66	3	17	2.23	(2)

Note: Entries in cells represent percentage of responses.

Table 1.8. Ranking of Personnel Types by Perceived Contribution to Public Risk

Factor	Rank 1	2	3	4	Mean Rank	Final Order
Mechanical technicians	21	29	39	11	2.40	(2)
Electrical technicians	4	43	39	14	2.63	(3)
I and C technicians	54	29	14	4	1.70	(1)
Other[a]						

[a]Insufficient data to calculate ranking.

Note: Entries in cells represent percentage of responses.

1.3. Summary and Discussion

While different points of view were sometimes expressed by the three groups of interviewees in response to specific questions included in the interview, all groups agreed on the overall usefulness of a structured computerized model for analyzing/predicting maintenance performance reliability, and the ability to provide information that would be applicable to risk and safety studies. No interviewee said that the type of information which a computer model could provide would be "slightly useful" or "not useful" and 80 percent replied "extremely useful" or "very useful."

The differences in point of view among the representatives of the various groups are not unusual for this type of work. In fact, some differences were anticipated at the outset and, accordingly, the various points-of-view analyses were instituted.

However, the scope of the interests of the various stakeholders suggests that the content of any computerized model must be quite flexible and include environmental, personnel, and motivational variables.

While a number of variables emerged as assuming highest usefulness ratings, none of the variables included in the interview was considered as "slightly useful" or "not useful" by any overriding proportion of the interviewees. Accordingly, any computerized model which lacks richness in content will, on the basis of the interview results, be considered barren by some stakeholder groups.

On the other hand, there is probably no limit to the number of variables and applications that are of interest. It is known that as models increase in generality, they decrease in validity. Also, more specific models are easier to conceptualize, structure, and develop. Accordingly, they are less costly to develop, and the time required for their development is less than for a more general representation. However, the applicability of highly specific models across a variety of situations is limited. As a result, the cost of ownership cannot be allocated across as many uses. While more general models are more costly in terms of developmental cost, such models are apt to be used regularly by a greater number of users for a greater number of purposes. For such more general models, the cost of development, maintenance, and support can be amortized over a number of uses and users and the total cost of owning one general model is almost certainly less than the cost of owning several more specific models.

In summary, the interviews pointed out a host of potential model variables and, to some extent, prioritized the variables. However, it should be pointed out that: (1) pseudorealism can be easily brought to an excess to the detriment of a model and of all concerned, and (2) the true test of a model is its ability to assist in decision making and not the number of relational equations which give the appearance of reflecting the real world.

2. MAIL SURVEY

2.1. Introduction

One purpose of the structured interviews was to serve as a test bed for a questionnaire to be mailed to a broader sample. The interviewees indicated few needs for extension/clarification of questions. They also seemed to think that the questions were clear and that the coverage was comprehensive. Accordingly, only a few changes were made to produce the mail questionnaire (Appendix C). The changes that were instituted included slight changes in wording and reorganization. That so few changes were required is indicative of the completeness of the initial interviews.

In its final form, the survey consisted of three parts. The questions in the first section asked for information about the respondent's background: experience, position title, position responsibilities, etc. The second section centered on the types of information a computerized model should provide. The possible types of information included the influence of such factors as team composition, team proficiency, training, heat, and stress on performance. Questions in the third section mainly dealt with human error. They asked about the types of tasks that are most likely to be performed incorrectly, the perceived public risk associated with various maintenance tasks, and the overall usefulness of a computerized model.

2.1.1. Questionnaire Mail Out

A larger number of the potential users of a structured computerized model was identified. Again, these users were representatives of nuclear power plants, architectural and engineering organizations, and the Nuclear Regulatory Commission. Copies of the final questionnaire, along with instructions for its completion and return, were mailed to these potential users. Approximately two weeks after the initial mail out, followup letters were sent to questionnaire recipients who had not returned their questionnaires by that time.

2.1.2. Questionnaire Return

Approximately five weeks after the initial questionnaire distribution, the returns were analyzed. A total of 160 surveys were distributed and, at the time of data analysis, 68 had been returned. This represents a return rate of 42.5 percent. The return rate by type of respondent is shown in Table 2.1. It is evident that the representatives of the Nuclear Regulatory Commission showed the greatest degree of cooperation with this work.

2.1.3. Description of Respondents

The first section of the mail survey requested actuarial information from the respondents. This information serves to provide insight into the background/experience of the respondents and, possibly, the scope of the information provided. The distribution and return

Table 2.1. **Mail Survey Distribution and Return**

Type of Organization	Original Distribution	No. of Returns	Return Percent	Percentage of Total
Nuclear Power Plant	108	42	38.89	61.76
Nuclear Regulatory Commission	12	7	58.33	10.29
Architectural/Engineering Organization	40	19	47.50	27.94
Total	160	68	42.50	100.00

rate[a] of the survey forms is given in Table 2.1. We note that, as in all surveys of this type, the respondents are self selected and do not necessarily represent a random sample of representatives of these three areas of concern in the nuclear power field. The original distribution lists were developed by the project staff who attempted to provide national coverage which was representative in a broad way, but not necessarily according to the formal standards of survey sampling theory. However, the overall return rate of 42.5 percent is quite reasonable and should afford confidence in the results.

The three interests surveyed were nuclear power generating stations (hereafter referred to as plants), the Nuclear Regulatory Commission (NRC), and architectural/engineering organizations. Respondents from plants made up 62 percent of the total return. The NRC respondents constituted 10 percent, and the architects/engineers made up 28 percent of the respondents.

The introductory section of the survey also inquired into the respondents' length of experience in the nuclear energy field. The response distribution is shown in Table 2.2.

As a group, the respondents appear to possess mid to long term experience with nuclear energy. This experience level was to be expected because the survey was intended for senior personnel.

Each respondent's position title was requested, and the types of titles reported are presented in Table 2.3. The high frequency of superintendents, section leaders, directors, and senior vice presidents in Table 2.3 clearly illustrates that the surveys were completed by senior personnel.

One further background informational question asked respondents about their area of responsibility. Table 2.4 lists both the categories named in the survey form as well as the areas supplied by the respondents in the "other" category.

---

[a]In tables to follow, the general convention will hold that when "all respondents" are referred to, N = 68; for the separate groups, generally "plants" N = 42, "NRC" N = 7, and "architect/engineering organization" N = 19. Although respondents occasionally left questions unanswered, the number of missing values was always small and never over five.

Table 2.2. Years of Experience in Nuclear Energy of Respondents

Years of Experience	Frequency	Percent
0 - 2	5	7.4
2 - 10	29	42.6
Over 10	34	50.0

Table 2.3. Position Titles of Respondents

Organization	Title	Frequency
Plant	Maintenance superintendent	7
	Training coordinator	7
	Supervisor of planning/scheduling	6
	Maintenance supervisor	5
	Foreman	3
	Associate engineer for training	2
	I & C supervisor	2
	Instrument engineer	2
	Mechanical maintenance engineer	2
	Technical assistant maintenance engineer	2
	Assistant superintendent, I & C	1
	Assistant superintendent, maintenance	1
	Branch engineer	1
	Electrical maintenance engineer	1
NRC	Section leader, procedures and test review	2
	Branch chief	1
	Nuclear engineer	1
	Reactor safety engineer	1
	Resident inspector	1
	Risk assessment engineer	1
Architect/ Engineer	Direction	3
	Senior vice president	3
	Chief, health physics	2
	Manager, I & C	2
	Senior scientist	2
	Engineering supervisor	1
	Human factors engineering group supervisor	1
	Manager, applications engineering and planning	1
	Marketing manager	1
	Repair superintendent	1
	Systems engineer	1
	Vice president, marketing	1

Table 2.4. Areas of Responsibility

Organization	Area	Frequency	Percent
All respondents	Mechanical maintenance	16	24
	Electrical maintenance	1	2
	Instrument and control maintenance	7	10
	Other	44	64
	Analysis of "Other" Responses		
Plant	Training	7	
	Mechanical, electrical, and I & C maintenance	6	
	Mechanical and electrical maintenance	4	
	Planning/scheduling	4	
NRC	Procedures in general	2	
	Human factors and risk	1	
	Initial testing and start up	1	
	Operator qualification and training	1	
	Risk analysis/accident sequence	1	
	Safety and quality assurance	1	
Architect Engineer	Health physics/radiation	4	
	Research and development	2	
	Applications engineering	2	
	Human factors engineering	2	
	Management	1	
	Marketing	1	
	Start up testing	1	
	Reliability and safety	1	

In summary, the respondents generally possessed depth of experience in the nuclear power field and held positions of senior responsibility in their organization. The areas of responsibility represented by the respondents were quite broad. Although the majority of the respondents were directly involved with maintenance, there was also representation in such areas as training, scheduling, risk analysis, human factors, health physics, reliability, management, and marketing. The respondents were clearly the type of potential user the survey was intended to reach. Their responses to the survey questions are described in the subsequent sections of this chapter. Following the organization of the survey form, the discussion is presented in two parts. First, the respondents' opinions are presented about the usefulness of various types of information that a computerized methodology pertaining to nuclear plant maintenance might provide. Then, the information provided by the respondents about factors influencing human error in nuclear plants is presented.

In treating the data, the responses from all three groups were combined. Differences among the groups will be pointed out only in those instances where they were marked and/or statistically significant.

2.2. Identification of User Needs

The mail survey was designed to assess the usefulness of various types of information which could be provided by a computerized model to potential users of the model. Quite obviously, such information possesses direct implications for the content of such a model. The majority of the questions asked the respondent to identify the most important factors influencing maintenance performance.

2.2.1. Usefulness of Various Types of Information

The survey contained 26 questions which asked the respondent to indicate the usefulness of specific information along a five category rating scale which ranged from "1" ("extremely useful") to "5" ("not useful"). In much of the discussion to follow these questions will be identified by a keyword or phrase. For example, "Noise" will be used to refer to the question on "How useful would information be about how noise level affects performance of a given maintenance action?" This was question 20 of the survey form. A list of the keywords and the associated question numbers is given in Appendix D. By consulting this appendix, the reader will be able to locate the original question as it was included in the survey.

Table 2.5 presents the results yielded by the "usefulness" items. The numbers represent the percentage of the total respondents who indicated that the information specified would be "extremely useful"; "very useful"; "useful"; "slightly useful"; or "not useful." One aspect of this table that is immediately apparent is the low percentages of responses in the "not useful" column. Almost one-third of these entries are zero. None of the percentages in the "not useful" column is over 10 percent. Accordingly, it appears that all of the factors affecting maintenance performance, as identified in the survey, are of some use to the survey respondents and, by implication, are of potential value for inclusion in a computerized model. The task now remains to determine which information the respondents would find most useful.

2.2.2. Prioritization

One indication of the highest priority information might be gained by summing the responses in the "extremely useful" and "very useful" categories. This procedure was completed, and the resulting percentages for the most highly rated information is given in Table 2.6.

The summary presented in Table 2.6 indicates the greatest need of the respondents to be information related to management decisions, i.e., information about how different levels of training and training emphasis affect maintenance performance, as well as information about how much time maintenance tasks require, who should be assigned to maintenance teams, and how the effectiveness of that team will vary with the proficiency of the team members. Two other types of information that were highly rated were concerned with

Table 2.5. Percentage of Responses to Usefulness Items

Items[a]	Extremely Useful 1	Very Useful 2	Useful 3	Slightly Useful 4	Not Useful 5
Cost-effective team	11.8	29.4	38.2	17.6	2.9
Exposure	26.5	35.3	25.0	8.8	4.4
Time required	23.5	36.8	27.9	11.8	0
Team size	27.9	29.4	36.8	5.9	0
Team composition	27.9	32.4	29.4	7.4	2.9
Time adjustment	17.6	25.0	33.8	23.5	0
Probability of error	27.9	27.9	25.0	14.7	4.4
Contracting out	17.6	36.8	29.4	11.8	4.4
Noise	5.9	16.2	45.6	26.5	5.9
Temperature	16.4	34.3	29.9	17.9	1.5
Gear	25.4	28.4	29.9	16.4	0
Tight places	9.0	31.3	31.3	23.9	4.5
Time constraints	16.4	29.9	40.3	13.4	0
Stress	14.9	41.8	23.8	10.4	0
Morale	13.2	32.4	32.4	17.6	4.4
Training	39.7	36.8	19.1	4.4	0
Proficiency	20.6	39.7	25.0	13.2	1.5
Time of day	10.3	20.6	36.8	26.5	5.9
Fatigue	25.0	26.5	30.9	16.2	1.5
Subtask success probability	13.4	35.8	37.3	11.9	1.5
Undetected errors	27.9	35.3	27.9	7.4	1.5
Risk	20.6	29.4	33.8	16.2	0
Demand vs. PM	14.9	22.4	28.4	25.4	9.0
Safety	25.0	32.4	32.4	8.8	1.5
Human reliability	17.6	35.3	30.9	19.7	1.5
Overall	13.2	41.2	22.1	19.1	4.4

[a]See Appendix D for key to original item (Appendix C).

Table 2.6. Percentage of Respondents Indicating Information Would be "Extremely Useful" or "Very Useful"

Information Type	Percent
Training	76.5
Undetected errors	63.2
Exposure	61.8
Time required	60.3
Team composition	60.3
Proficiency	60.3

health physics and quality assurance: information on radiation exposure levels and information on undetected errors. A comment added by one of the plant respondents is useful here in terms of a summary and illustration:

> *The information you asked about would be useful because now these decisions are based on personnel [sic] experience (which can never be replaced in an individual situation). Anything to aid in making these decisions in a more scientific manner would be useful.*

Although the preceeding discussion provides an indication of the relative importance or usefulness of the 26 different types of information covered in the survey, the five separate categories of usefulness make these comparisons somewhat difficult. A single measure of an item's importance would alleviate this problem. Accordingly, the mean usefulness of each item was calculated with the numerical values of 1 through 5 representing the categories of "extremely useful" through "not useful." The overall mean usefulness ratings are presented in Table 2.7. This table gives in a single measure the relative importance of the various types of information. By this measure, the overall most important item was concerned with training (mean = 1.88), while the item with the lowest rating dealt with the effects of noise on performance (mean = 3.10).

Table 2.8 lists the five items which had the highest usefulness ratings, along with the five items with the lowest usefulness ratings. It should be recalled that even for items with low ratings, overall, each of these items was rated as at least "slightly useful" by at least 90 percent of the respondents. The lowest rated items are identified not to imply that they are unimportant, but rather to make clear that other information was viewed as more important. The implication is that if limited resources are available for developing a model of maintenance reliability, variables with a higher importance or usefulness rating should be included in the model before those variables rated as less important.

As would be expected, the types of information identified in Table 2.8 as having the overall highest usefulness rating are in close agreement with the most useful information given in Table 2.6. Once again, "training" is the category with the highest usefulness rating. The only type of information given in Table 2.8 that was not included in Table 2.6 is "number." The implication for a computerized model is clear: a model must include the factors of training, undetected errors, exposure, time required for task completion, team composition, team size, and the proficiency of team members.

2.2.3. Points of View Analysis

Representatives of the NRC and of architectural/engineering organizations may have needs and priorities which are different from those of persons who work with plant maintenance. Accordingly, the responses to each of the questions listed in Table 2.5 were further categorized by the type of respondent: (a) plant, (b) NRC, or (c) architectural/engineering. Because of the limited number of NRC respondents, the response categories were collapsed, combining the "extremely useful" category with the "very useful" category, and the "slightly useful" category with the "not useful" category. This resulted in a set of 3 x 3 contingency tables and, while not eliminating the statistical problems, greatly reduced their impact.

Table 2.7. Mean Usefulness Rating

Item	Rating
Cost-effectiveness team	2.71
Exposure	2.29
Time required	2.28
Team size	2.21
Team composition	2.25
Time adjustment	2.63
Probability of error	2.40
Contracting out	2.48
Noise	3.10
Temperature	2.54
Gear	2.37
Tight places	2.84
Time constraints	2.51
Stress	2.39
Morale	2.68
Training	1.88
Proficiency	2.35
Time of day	2.97
Fatigue	2.43
Subtask success probability	2.52
Undetected errors	2.19
Risk	2.46
Demand vs. PM	2.91
Safety	2.29
Human reliability	2.47
Overall	2.60

Table 2.8. Mean Usefulness Ratings Over All Respondents

	Item	Rating
Highest:	Training	1.88
	Undetected errors	2.19
	Team size	2.21
	Team composition	2.25
	Time required	2.28
Lowest:	Cost effective team	2.71
	Tight places	2.84
	Demand vs. PM	2.91
	Time of day	2.97
	Noise	3.10

For several items for which statistically significant differences were noted among the groups, there was a general pattern. Two of the three groups were quite close in their rating of the usefulness of some specific information and one group showed a widely dissimilar pattern. In four of the six items to be discussed, the plant respondents and architects/engineers demonstrated similar response patterns. For two questions, the NRC respondents and architects/engineers showed similar patterns. A clear example of this grouping is seen in the item which sought information about the usefulness of information on the relationship between maintenance team proficiency and the probability of errors made during maintenance. The results are shown in Table 2.9.

There are statistically significant differences, $\chi^2(4) = 9.88$, $p < .05$, among the groups. One hundred percent of the NRC respondents said that error probability information would be "extremely useful" or "very useful." The replies of respondents from plants and architectural/engineering organizations were more equivocal.

Tables 2.10, 2.11, and 2.12 also show a unique response pattern by the NRC respondents. Information about task completion time, required number personnel, and the errors associated with preventive as compared with demand maintenance was rated as more useful by plant and architect/engineer respondents.

Table 2.10 presents the data on how useful information would be about the time required for completing maintenance tasks. The differences shown in Table 2.10 are statistically significant at or below the .05 level of confidence, $\chi^2(4) = 12.01$, $p < .05$. Perhaps this response pattern resulted because information of this type was perceived to be more relevant to the day-to-day maintenance and service of a plant (concerns of plant and architect/engineer respondents) and not immediately relevant to the overall safe operation of a plant (an NRC concern). On the other hand, down time for a safety related system would certainly seem to be an NRC concern. The conclusion relative to the differential perception is supported by the results shown in Tables 2.11 and 2.12, which summarize the results from items inquiring into: (1) the number of technicians to be assigned to a maintenance task, $\chi^2(4) = 9.51$, $p < .05$, and (2) the likelihood of error during demand maintenance as compared with preventive maintenance, $\chi^2(4) = 9.23$, $p < .05$. The response percentage to both of these questions (plant and architect > NRC) was depressed for NRC respondents, and both questions were concerned with daily plant operations.

There were two questions for which the respondents from plants stood apart from the NRC and architect/engineer respondents. Table 2.13 presents the response distribution for the question dealing with temperature effects. Plant respondents were more likely to rate this information "extremely useful" or "very useful" than were the NRC or the architect/engineer respondents, $\chi^2(4) = 11.44$, $p < .05$.

Differences among these three groups were also noted, although this difference did not reach statistical significance, for information about maintenance team composition while keeping exposure levels within bounds. Such information was judged more useful by the NRC and the architect/engineer respondents, $\chi^2(4) = 6.97$, $p = .14$. The response distribution relative to this issue is shown in Table 2.14.

Table 2.9. Usefulness Rating by Group Membership — Probability of Error

Category	Plant	NRC	Architect/ Engineer
Extremely/Very Useful	43*	100	68
Useful	31	0	22
Slightly/Not Useful	26	0	10

Table 2.10. Usefulness Rating by Group Membership — Time Required

Category	Plant	NRC	Architect/ Engineer
Extremely/Very Useful	60*	14	79
Useful	24	72	21
Slightly/Not Useful	16	14	0

Table 2.11. Usefulness Rating by Group Membership — Number of Technicians Needed

Category	Plant	NRC	Architect/ Engineer
Extremely/Very Useful	55*	14	79
Useful	38	72	21
Slightly/Not Useful	7	14	0

Table 2.12. Usefulness Rating by Group Membership — Demand vs. PM

Category	Plant	NRC	Architect/ Engineer
Extremely/Very Useful	33*	14	56
Useful	29	14	33
Slightly/Not Useful	38	71	11

*Cell entries represent percentage within each category.

Table 2.13. Usefulness Rating by Group Membership — Temperature

Category	Plant	NRC	Architect/ Engineer
Extremely/Very Useful	60*	29	39
Useful	16	71	44
Slightly/Not Useful	24	0	17

*Cell entries represent percentage within each category.

Table 2.14. Usefulness Rating by Group Membership — Exposure

Category	Plant	NRC	Architect/ Engineer
Extremely/Very Useful	50*	72	84
Useful	33	14	10
Slightly/Not Useful	17	14	6

*Cell entries represent percentage within each category.

Perhaps the plant respondents thought that presently existing methods are adequate for predicting radiation levels and exposure. However, we note that half of the plant respondents indicated that this information would be "extremely useful" or "very useful." This value is modest only in comparison with the very large percentages of NRC and architect/engineer respondents who indicated that this information would be "extremely useful" or "very useful."

2.2.4. Means Analysis

The preceding analyses were based on group differences in response distribution. An alternate approach to investigating group differences is through an analysis of means. This approach was taken earlier when the items were prioritized. The mean usefulness ratings for each of the three groups are given for the highest rated and lowest items in Table 2.15.

Table 2.15 again suggests some lack of concordance among the three groups in the ranking of the usefulness of information items. For example, "training" was highly ranked by the plant and the NRC respondents but not so ranked by the architects/engineers; the plant and architect/engineer respondents agreed that information about the number of technicians in a team is highly important, but the NRC respondents did not rate this information so highly. On the other hand, all three groups included information about noise effects as of relatively low usefulness.

Table 2.15. Mean Usefulness Rating* by Respondent Type

	Plant		NRC		Architect Engineer	
	Item	Mean Rating	Item	Mean Rating	Item	Mean Rating
Highest:	Training	1.81	Probability of error	1.86	Team size	1.68
	Safety	2.19	Training	2.00	Team composition	1.74
	Team size	2.31	Undetected errors	2.00	Time required	1.79
	Fatigue	2.33	Proficiency	2.14	Exposure	1.80
	Undetected errors	2.33	Gear	2.29	Undetected errors	1.95
			Human reliabiliy	2.29		
Lowest:					Temperature	2.67
	Time adjustment	2.83	Cost-effective	3.14	Tight places	2.67
	Tight places	2.86	Noise	3.14	Overall	2.68
	Demand vs. PM	3.05	Tight places	3.14	Noise	2.79
	Time of day	3.12	Morale	3.14	Morale	2.79
	Noise	3.24	Demand vs. PM	3.71	Time of day	2.84

*1 = "extremely useful;" ... 5 = "not useful."

2.2.5. Profile Similarity

To gain an indication of the overall similarity in ranking of information usefulness among the groups, Spearman rank order correlation coefficients were computed. The plant and architect/engineer respondents showed moderate, but statistically significant agreement in their rankings, $\rho = .46$, $p < .29$. The correlations between plant and between NRC respondents and NRC and architect/engineer respondents were lower and were not statistically significant, $\rho = .29$ for both. These correlation coefficients apparently reflect the similarity in perspective toward maintenance of the plant and architect/engineer respondents. As suggested previously, this pattern may reflect the different goals of the three organizations. The NRC is charged with the responsibility of plant and public safety. The plant and architect/engineer respondents are more concerned with the daily plant activities, maintenance, and service.

Figure 2.1 illustrates the mean usefulness ratings for all three groups over all 26 items. It is apparent, from Fig. 2.1 that the groups' responses were very similar on some items (e.g., stress) and quite disparate on other items (e.g., time required, and number).

Analyses of variance were performed on the data summarized in Fig. 2.1. Several of the group differences were found to be statistically significant.

The mean ratings on the item inquiring about the usefulness of information relative to the amount of time required by maintenance tasks were significantly different, $F(2,65) = 5.28$, $p < .01$. Subsequent *t* tests revealed that the architects/engineers rated this item more highly than either the plant or the NRC respondents. The group differences on the "number" items were statistically significant, $F(2,65) = 6.93$, $p < .01$, as were those on the

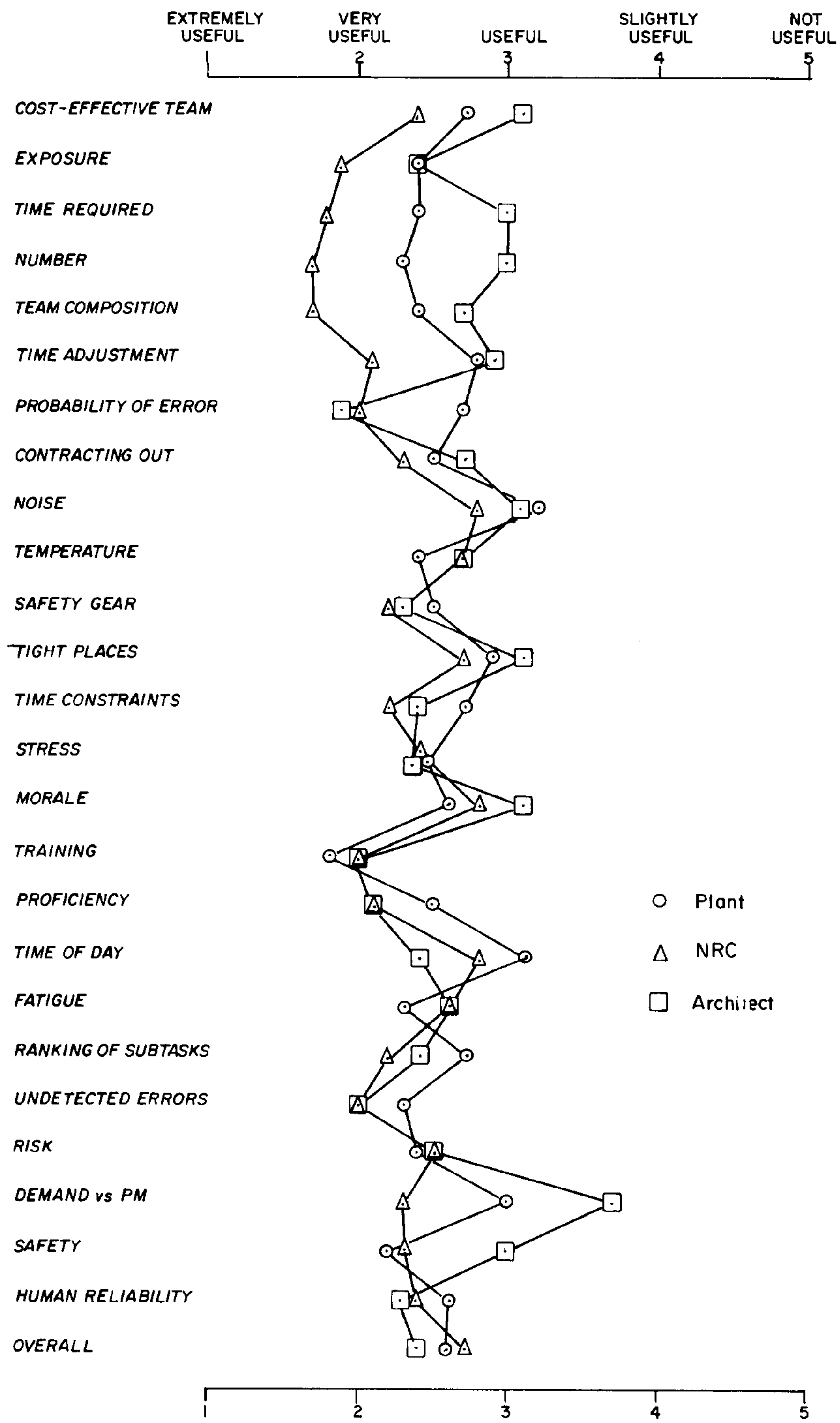

Fig. 2.1. Mean Usefulness Ratings by Respondent Group.

"team composition" item, $F(2,65) = 3.75$, $p < .05$. For both of these items, the architect/engineer ratings were significantly higher than the plant or NRC ratings. The "time constraints" item showed significant differences, $F(2,65) = 3.69$, $p < .05$. On this item, *t* tests indicated that only the difference between plant and architect/engineer ratings was reliable.

The "demand vs. PM" item was the final question for which a statistically significant difference was indicated, $F(2,64) = 4.83$, $p < .05$. Again, *t* tests indicated that the architect/engineer rating was significantly higher than either the plant or NRC ratings.

The ratings on the "probability of error" item approached statistical significance, $F(2,65) = 3.12$, $p < .06$. This trend appears to have resulted from the relatively lower plant ratings.

In general, the pattern of differences was that the architect/engineer respondents rated the information dealt with in these items as more useful than the plant or NRC respondents. The reason for this overall more positive view by the architect/engineer respondents is not immediately apparent. One conjecture explaining the more positive view taken by the architect/engineer respondents is that they are more accepting of computer technology in general.

2.2.6. Direct Comparisons

The preceding discussion described the relative importance of different types of information by comparing their separate usefulness ratings. The survey contained two items in which the respondent was asked to compare directly various types of information and to rank order the information types in terms of relative usefulness. The results from one set of these items are shown in Table 2.16, which presents the percentage of all respondents who, when ranking each item, gave that item a first, second, or third, etc., ranking. For example, 3 percent of the respondents ranked information on noise in first place. Table 2.16 additionally gives the mean ranking of each item and, in parentheses, the overall ranking of means. Of the information categories included, the respondents thought that information about the effects of wearing safety gear was more important than the other types of information with which it was compared.

Table 2.17 presents the percentage of respondents in each group who gave first place ratings to the various types of information. Table 2.17 indicates that information about the effects of wearing safety gear was most frequently ranked first by only the NRC respondents. Information about the effects of time constraints was more frequently ranked first by plant and architect/engineer respondents. "Time constraints" was also ranked first by a large percentage of the NRC respondents. In summary, it seems that information about the effects of wearing safety gear and about the presence of time constraints was regarded as more important by all respondents than information on the effects of heat, noise, or tight working conditions.

Table 2.18 presents the results of the ranking of four types of performance moderating variables. As in Table 2.16, the entries represent the percentage of all respondents who gave a particular ranking to the specified information. Of the factors listed, information

Table 2.16. Ranking of Usefulness of Selected Working Conditions Information Types

Item	Rank 1	2	3	4	5	Mean Rank	Final Order
Noise	3	9	12	22	53	4.125	(5)
Heat	21	33	9	33	3	2.636	(3)
Gear	32	31	23	8	6	2.246	(1)
Tight places	5	16	31	30	19	3.422	(4)
Time constraints	42	15	19	9	15	2.403	(2)

Table 2.17. Percentage of First Place Ranking — Working Conditions

Item	Plant	NRC	Architect/ Engineer
Noise	2	0	5
Heat	29	0	10
Gear	27	57	30
Tight places	5	0	5
Time constraints	37	43	50

Table 2.18. Ranking of Information Usefulness — Performance Moderating Variables

Item	Rank 1	2	3	4	Mean Rank	Final Order
Stress	18	15	32	35	2.833	(4)
Training	47	29	15	9	1.853	(1)
Morale	12	26	30	32	2.818	(3)
Proficiency	26	33	20	21	2.364	(2)

about the effects of training was clearly the most highly regarded. Investigation of the ranking patterns of the three respondent groups indicated that "training" was consistently ranked in the first place by all three groups. This result further supports the general high evaluation of information on the effects of training. The consistency of this finding indicates the high esteem in which training is held throughout the nuclear industry. Training is emphasized for productivity as well as for safety reasons. It is clear that, just as training holds a central part in the nuclear industry, it must play an important role in any type of methodology addressing the maintenance context that hopes to provide benefit to the nuclear industry.

2.2.7. Other Personnel Factors

Other personnel variables are important to overall plant safety, efficiency, and reliability. The survey also asked respondents to indicate the factors, in terms of personal characteristics, employment history, etc., which are important in the selection of maintenance technicians. The responses are summarized in Table 2.19. The respondents were instructed to indicate from a menu the two most important items, and the entries in Table 2.19 represent the percentage of all responses to each factor. The most frequently chosen factors were "positive work attitude" and "precision in performing a task."

As might be expected, the three groups were not entirely congruent in their selection of personnel factors. For example, "positive work attitude" was chosen by 57 percent of the plant respondents, as contrasted with 43 percent of the NRC respondents and 37 percent of the architect/engineer respondents. On the other hand, "prior experience" was regarded as more important by the NRC and architect/engineer respondents (43 percent and 42 percent, respectively) than by the plant respondents (33 percent). Finally, "mechanical ability" was chosen by 38 percent of the plant respondents, 26 percent of the architect/engineer respondents, and zero percent of the NRC respondents.

Table 2.19. Percentage of Responses to Various Personnel Factors

Item	Percentage of Responses
Intelligence	13
Mechanical ability	15
Determination to get things done	6
Positive work attitude	25
Prior experience	18
Precision (accuracy) in performing a task	20
Others	2

2.2.8. Level of Model Detail

Two questions in the survey asked for information about the nature of the output that should be provided by a computerized technique. Fifty-two percent of the respondents indicated that they would want information at both the subtask level and at the total task level. This compares with 28 percent of the respondents who preferred the information only at the subtask level and 20 percent of the respondents who wanted information only at the task level. It appears that users of a computerized model prefer to have both levels of analysis available so that either level could be consulted depending on the specific circumstances involved.

Finally, in terms of the nature of the computer output, 47 percent of the respondents indicated that they preferred to have the output information available in a hard copy printout form as compared with a cathode ray tube (CRT) display. Twenty-one percent preferred the CRT display, while 32 percent responded that the output should be available in both forms.

2.3. Causes and Consequences of Maintenance Errors

The third section of the survey dealt with the nature and effects of maintenance errors. Questions concerning the relative frequency of errors, the environmental and personal factors contributing to errors, and the consequences, in terms of perceived public risk, resulting from errors were included. Such information provides a partial basis for enhancing the content validity of any simulation model that might emerge from this work.

2.3.1. Types of Maintenance Error

The respondents were asked about the types of maintenance tasks that are most likely to be performed incorrectly in nuclear power plants. The results are summarized in Table 2.20. In contrast with the differences in opinion among the three interest groups shown on this question during the structured interviews, the three respondent groups to the mail survey were fairly homogeneous in their responses. "Fault diagnosis" was the most frequent response (76 percent of plant personel, 43 percent of NRC personnel, 56 percent of architect/engineer personnel). Despite the agreement on the most frequent type of error, the groups again differed on the likelihood of preventive maintenance being incorrectly performed. For the architect/engineer respondents, preventive maintenance was the second most frequently selected alternative (50 percent). However, respondents representing plants and the NRC were not as likely to reply that preventive maintenance is often incorrectly performed (21 percent of plant respondents, 29 percent of NRC respondents indicated this alternative).

2.3.2. Causes of Error

The survey attempted to determine not only the relative frequency of error but also the causes of error. Two questions asked the respondents to rank various factors in accordance with their contribution to error. The resultant ranking of the relative contribution of environmental factors is given in Table 2.21. As with the previous tables of this type, the cell entries represent the percentage of respondents who gave a particular rank to each listed factor. Table 2.21 indicates that, overall, "time constraints" and "heat" were said to be the greatest contributors to the possibility of human error during maintenance.

The three respondent groups were in general agreement on error cause, just as they were in identifying the most likely error. "Time constraints" was ranked first by 48 percent of plant respondents, by 57 percent of NRC respondents, and by 57 percent of architect/engineer respondents. This represents a strong consensus on the importance of the presence of time constraints to understanding (in terms of prediction and control) errors. While this consensus existed for the primary factor, no such agreement existed for

Table 2.20. Most Probable Errors in Nuclear Power Plant Maintenance

Type of Error	Percentage of Responses
Fault diagnosis	34
Repair	19
Replacement	6
Checkout	9
Calibration	8
Preventive maintenance	14
Inspection	7
Other	2

Table 2.21. Ranking of Factors by Contribution to Maintenance Error

Factor	Rank 1	2	3	4	5	Mean Rank	Final Order
Noise	3*	1	11	30	54	4.318	(5)
Heat	22	30	22	18	8	2.582	(2)
Safety gear	21	19	16	25	18	3.000	(4)
Time constraints	53	13	19	9	6	2.015	(1)
Tight places	4	39	29	15	12	2.909	(3)

*Cell entries represent percentage of responses.

Table 2.22. Percentage of Respondents Giving First Ranking to Maintenance Error Inducing Factors

Factor	Plant	NRC	Architect/ Engineer
Noise	0	14	5
Heat	29	0	14
Safety gear	19	29	19
Time constraints	48	57	57
Tight places	5	0	5

the relative importance of the other factors. For example, "heat" was ranked first most frequently after "time constraints" by plant respondents (29 percent), while zero percent of the NRC respondents and 14 percent of the architect/engineer respondents ranked "heat" first. These results are shown in Table 2.22, which gives the percentage of respondents giving first ranking to each of the factors. There is a discrepancy in the ranking not only of "heat" but also of "noise." Perhaps, the "heat" item was ranked as relatively more important by the plant respondents because they must deal with it most directly or must deal with maintenance technician complaints about it most often. Table 2.22 indicates that there was relative agreement among the three groups on the contributions of "safety gear" and "tight places" to error.

2.3.3. Undetected Error

The respondents also ranked a number of factors with regard to their contribution to the likelihood of undetected errors. The results of this ranking are summarized in Table 2.23. The results indicate that, although "lack of experience" was the item most frequently ranked first, "not following procedures" was generally regarded to be the most important contributor to undetected error.

While the responses for the total group resulted in a final overall first ranking of "not following procedures," the individual respondent groups did not universally rank "not following procedures" in first place. This is illustrated in Table 2.24, which shows the percentage of respondents giving a first place rank to various factors. It can be seen that "lack of experience" was clearly considered by the plant respondents to be the major contributor to undetected errors. For the NRC respondents, "lack of experience" was also ranked first frequently, but just as frequently ranked first were the "not following procedures" and "poorly documented procedures" items. The architect/engineer respondents ranked first "poorly documented procedures" and "inadequate supervision" (tied).

2.3.4. Preventive Maintenance

In order to collect information about the nature of maintenance tasks in nuclear power plants, the survey included one question which inquired into the percentage of maintenance tasks which are preventive in nature. The resultant response distribution, for all respondents and for the separate respondent group, are shown in Table 2.25. The table indicates a difference of opinion between plant and architect/engineer respondents on this issue. The model category for plant respondents was "20-39 percent" (whereas architect/engineer respondents said that preventive maintenance is a larger percentage of maintenance tasks). Their model category was "60-79 percent." It is noted in this context that architect/engineer respondents were more likely than plant respondents to indicate that preventive maintenance is the area in which errors are likely to be made (see earlier discussion concerning Table 2.20). Finally, Table 2.25 indicates that there was no consensus of opinion among NRC respondents.

Table 2.23. Ranking of Factors by Contribution to Undetected Error

Factor	Rank 1	2	3	4	5	6	Mean Rank	Final Order
Poor system design	15*	8	18	9	21	29	4.000	(5)
Inappropriate tools	0	5	6	9	28	52	5.169	(6)
Not following procedures	27	42	18	18	9	3	2.667	(1)
Poorly documented procedures	17	32	20	18	9	3	2.800	(2.5)
Inadequate supervision	15	20	21	18	17	9	3.288	(4)
Lack of experience	31	15	14	25	14	1	2.800	(2.5)

*Cell entries represent percentage of responses.

Table 2.24. Percentage of Respondents Giving First Ranking to Undetected Error Contributors

Factor	Plant	NRC	Architect/ Engineer
Poor system design	16	0	14
Inappropriate tools	5	0	5
Not following procedures	27	29	19
Poorly documented procedures	9	29	24
Inadequate supervision	9	14	24
Lack of experience	34	29	14

Table 2.25. Preventive Maintenance as Percentage of Total Maintenance

Percentage	All Respondents	Plant	NRC	Architect/ Engineer
80-100	3*	0	0	13
60-79	27	21	29	40
40-59	20	21	29	13
20-39	33	41	29	13
0-19	17	17	14	20

*Cell entries represent percentage of responses.

2.3.5. Perceived Error Consequences

The third section of the survey also inquired into the perceived consequences of errors. One item asked what percentage of maintenance tasks, if performed incorrectly, would lead to an event report, site emergency, or general emergency. The responses to this question are summarized in Table 2.26. An analysis of the response patterns by respondent type indicated statistically significant differences among the groups, $\chi^2(8) = 15.79$, $p < .05$.

The NRC respondents indicated that negative consequences would result from errors made on a higher percentage of maintenance tasks than did plant and architect/engineer respondents.

In a question further investigating the causes and consequence of maintenance errors, the respondents were asked to rank order maintenance personnel types in regard to how much an error committed by them would potentially contribute to public risk. The responses are summarized in Table 2.27.

The majority of the respondents perceived that errors made by I&C technicians are potentially the most likely to involve some element of public risk. This overall pattern was also evidenced in the rankings of the three groups. I&C technicians were ranked first by 75 percent of plant respondents, 71 percent of NRC respondents, and 50 percent of the architect/engineer respondents. In terms of overall ranking, mechanical technicians, electrical technicians, and quality assurance inspectors followed I&C technicians in that order.

Table 2.26. Consequences of Maintenance Errors

Percentage	All Respondents	Plant	NRC	Architect/ Engineer
80-100	2	0	17	0
60-79	6	7	0	6
40-59	19	17	50	12
20-39	28	26	17	38
0-19	50	50	17	44

Table 2.27. Potential Contribution to Public Risk by Personnel Types

Type	Rank 1	2	3	4	Mean Rank	Final Order
Mechanical technicians	16*	33	43	8	2.429	(2)
Electrical technicians	6	41	41	12	2.594	(3)
I&C technicians	69	23	6	2	1.406	(1)
QA inspectors	11	8	8	73	3.429	(4)

*Cell entries represent percentage of responses.

2.4. Discussion and Summary

The respondents to the mail survey, as well as the interviewees, were persons with wide ranging interests and responsibilities in the nuclear power field. Based on their years of experience and position titles (e.g., maintenance superintendent, coordinator of training, director of planning and scheduling, section chief, director of research and development), these persons were able to provide informed opinions about the items of present concern.

Information on the effects of training vis-a-vis job proficiency, probability of error, time requirements, etc., was viewed as the most useful type of information a computerized model could provide. During the interview phase, 80 percent of the interviewees rated information on training as "extremely useful" or "very useful." This compares with 76.5 percent of the mail survey respondents. This consistency in results provided by two different methods increases confidence in the results of the total work. In addition to information on training, the two data acquisition techniques indicated that information on undetected errors, radiation exposure, maintenance technician safety, time required for task completion, team composition, and team proficiency is of the highest saliency. Although these types of information received the highest ratings, it is important to remember that all the respondents found all the various types of information offered in the survey to be useful to a greater or lesser degree. In the mail survey, no more than 10 percent of the respondents ever indicated that information of a particular type would be "not useful."

These findings suggest strong support for the utility and benefits of a computerized methodology that addresses maintenance activities and the information that such a model would provide.

Given that the need for a computerized model which addresses the acts and behaviors of maintenance personnel is established by the present data and that various users would benefit from the information provided by such a model, the questions of what type of methodology should this model be, and can the model be developed within the framework of the chosen methodology, must be addressed.

3. LITERATURE REVIEW OF HUMAN BEHAVIORAL METHODOLOGIES

3.1. Introduction

Previous sections of this report examined the front-end user requirements of a proposed computerized methodology that would address maintenance activities within the nuclear power plant (NPP) context. One of the questions raised was, what type of methodology should be developed? In order to answer this question, a literature review of existing human behavioral methodologies was performed.

The literature review that was accomplished was done in two steps. Initially, a general literature review was performed which identified major types of methodologies and techniques. Those that were felt to have little applicability or adaptability potential were eliminated as candidates and will not be discussed within this report. The methodologies that were not eliminated after Step 1 were investigated further in Step 2 and the results of this investigation are presented herein.

It is the purpose of this review to examine existing whole task methodologies for input into determining the type of methodology to be developed within this program. Pros and cons of each of the techniques examined are given, but comparisons of their effectiveness on similar tasks is beyond the scope of this review and were not addressed.

3.2. Background

Technological advances experienced during this century have effectively minimized the role of the human in complex, sophisticated, man-machine systems. Various safety, operational, economic and sociological concerns, however, have prevented the total disappearance of the versatile human element, especially within systems that have associated with them a potential for high public risk in the event of their failure. Although the human element has been retained, technological advances, up to about twenty years ago, were primarily with system hardware, with little emphasis on the relationship between hardware and the men expected to operate and maintain these systems. The last twenty years has produced a dramatic increase in the amount of concern over the relationship of man within man-machine systems (MMS).

In examining the historic perspective of man's role with respect to MMS, it can be seen that the nuclear industry, when compared to other highly technical industries, is a relative newcomer on the scene. Its advent and subsequently intense development, however, came at a time when primary emphasis was still placed on the design of system hardware with secondary considerations given to how people were to operate and maintain these complex systems. Essentially all of today's nuclear power plants (NPPs) were designed with considerations prioritized in this way. For a more detailed historic perspective of man's role with respect to man-machine systems, refer to Price, Maisano, and Van Cott.[6]

3.3. Human Reliability in System Phases

System reliability must take into account all phases of system development. Meister[7] notes that there are four principal stages of system development in which human errors affect system reliability. These are: design, production, testing, and operation. Under the heading of testing, Meister includes installation and maintenance errors. Of these four stages, operation and testing (maintenance) seem to be the principal areas whereby human errors contribute to system failure. Rigby[8] indicates that maintenance/installation errors and operator errors are the primary contributors to system failures during a systems representative life cycle. Meister[9] states that 42.9% of all human initiated failures reported for the Operational Systems Test Facility-1 (Vandenberg Air Force Base) for a twelve-month period were due to assembly and installation errors. An informal examination of Licencee Event Reports (LERs) done at ORNL for this program indicated that approximately 20% of all failures to safety systems for Light Water Reactors (LWRs) can be linked to errors in maintenance. It is clear that human errors committed during operations and maintenance have a distinct effect on system reliability.

Activities during operations and maintenance have the potential for causing system failures that have associated with them a high risk to public safety. The operations function, however, is viewed as having a much more immediate and active potential for causing high risk system failures when compared to the maintenance function. This viewpoint is reflected in the fact that much more human factors work has been centered in the operations function than any other phase of a system. Human factors work in operations has resulted in the design of safety oriented, system based, functional interlocks where high risk operator error potential has been explicitly identified. It has provided input for optimizing layouts of controls and indicators to minimize the potential for operator error. It has also led to the development of predictive operator reliability techniques, some of which have been validated. In comparison, the maintenance function has received considerably less attention. Primarily the work in this area has taken the form of human factor reviews of plant designs in order to ascertain the requirements for conducting effective maintenance. A recent example is an EPRI report[10] on a human factors review of five nuclear plants and four fossil fuel plants. Reports such as these are valuable in identifying areas of maintenance which have a potential, when changed, to increase maintenance personnel effectiveness. However, little emphasis, as yet, has been placed on the quantitative assessment of human reliability within the maintenance context. In particular, relatively little work seems to have been done on developing quantitative human reliability models, geared specifically for the maintenance of a system.

3.4. The NPP Maintenance Personnel Reliability Modeling Program — Primary Goal

The primary goal of the nuclear power plant maintenance personnel reliability modeling program is to develop a quantitative, validated, predictive methodology which can be used to evaluate the contribution of maintenance error to overall NPP risk. The format and structure of the methodology to accomplish this task should be tailored to be fully responsive to the needs and requirements of its ultimate users, and should be based on an assessment of existing methodologies. To these ends, a front-end user survey (discussed earlier in this report) and a literature survey of existing methodologies were completed. The results of the literature survey is the focus of this section of the report.

3.5. Literature Survey — General Aspects

A general literature review of existing human behaviorial methodologies has revealed that a large number of methodologies of various types exist. Many of these methodologies are highly specific in nature and have little if any applicability or adaptability potential for the maintenance context of system operations, e.g., models discussed by Davis,[11] Kramer,[12] and Penrod.[13] The review also indicated that several distinctively different types of methodologies have been developed. The types of methodologies, examples of specific models and an assessment of their applicability in the maintenance context are detailed later in this report.

The literature, and personal discussions with Embrey,[14] Swain,[15] and Meister,[16] reveals that essentially no new complete human reliability prediction techniques have been developed since about the mid-1970's. Although some of the techniques from this era have been refined, improved upon, and broadened, their underlying approach to the prediction of human reliability remains the same. Because of this, earlier reviews of human reliability models by Meister[17,18] and Embrey[19] were used to gain insight into the major models which may have applicability in the maintenance area.

3.6. Quantitative Techniques — An Overview

As mentioned earlier, with respect to the operations context, relatively little work has been done with respect to the development of predictive human reliability models specifically for the maintenance context of system operations. In fact, the majority of the existing, major human reliability techniques have been developed primarily for system operations. Embrey[19] states that, in theory, any of the predictive techniques applied in the operational context should also be usable to analyze maintenance tasks. He states further, however, that the application of operationally based techniques to the maintenance context may pose problems due to the inherently different nature of maintenance tasks as compared to operational tasks. Table 3.1 lists some of these differences. Because of such differences, the direct application of existing human reliability prediction techniques to the maintenance area has been generally unsuccessful. The techniques do, however, represent a significant effort at quantifying human performance and were therefore examined in order to determine their applicability in the maintenance context.

Table 3.1. Some Characteristics of Maintenance Tasks As Compared to Operation Tasks

1. Maintenance tasks are more labor intensive.

2. Consequences of maintenance errors are more difficult to anticipate.

3. Many errors committed during maintenance activities do not become immediately apparent.

4. Error prevention due to appropriate hardware interlocks is not a prevalent in the maintenance area.

5. Maintenance task performance is generally less structured.

6. Maintenance tasks, especially during troubleshooting, are more cognitive in nature.

3.7. Desirable Qualities of a Quantitative, Predictive Technique

Before examining the various, complete techniques for the prediction of human reliability, some desirable qualities of a predictive technique are presented. This list of qualities, although not an exhaustive list, summarizes characteristics that were identified through preliminary visits to NPPs and were felt to be important and/or desirable within a quantitative predictive technique. These qualities were kept in mind as the various techniques were reviewed and aided in the determination of the type of methodology to be developed for this program.

1. The technique should be general in nature. The methodology should be able to handle any maintenance position, and should be able to generate results for teams of individuals with various maintenance titles and levels.

2. The technique should be able to easily handle interdependencies between task elements. The human in a man-machine situation has been described[19] as a multivariate, non-linear, reactive system component whose behavior is inherently statistical in nature. From practical experience, man has been shown to be prone to "common-mode" type failures. It is thus without a doubt that behavioral task elements tend to be interdependent and the development of any technique for human reliability prediction should be sympathetic to this fact.

3. The technique should utilize both human and equipment characteristics. Interdependencies between behavioral task elements in a man-machine situation stem from both human characteristics and equipment characteristics. The human factors work in the maintenance area to date has emphasized the design and layout of hardware for optimal human interaction. Equal consideration must be given to inherent human characteristics such as stress motivation, etc., which also have an effect on the operation and maintenance of systems and their hardware.

4. The technique should supply outputs of many variables of interest to the user. Any technique for human reliability prediction should supply to its user a summary of values of important variables of interest. A spectrum of values of variables will allow the user to gain more of an insight into the details of the task being analyzed, more so than the generation of a single global value such as probability of successful completion, or overall system reliability. For example, an analyst who is supplied with the probability of success of all of the subtasks within a task can much more readily identify problem areas with the task, whereas such identification is not easily achieved when the overall task probability of success is the only output parameter generated. In addition, the time variation of parameters such as stress, fatique, etc., are also desirable for problem area identification.

5. The technique should include capabilities for handling cognitive task elements. Human reliability prediction techniques to date have lacked the capability for handling task elements that are highly cognitive in nature. Although problem-solving theories, such as the Simon-Newell Theory[20,21] have led to the development of decision making modules,[22] their application within complete human reliability prediction techniques has been rather limited. Because the maintenance area includes many tasks that are inherently cognitive in nature (especially during troubleshooting), the incorporation of decision making routines or modules should be a significant concern in the development of any human reliability prediction technique for maintenance activities.

6. The technique should provide means for sensitivity analysis of parameters within the methodology. In order to assess the impact of proposed changes to system design, maintenance procedures, training, team structure, etc., it is desirable to have a technique which will easily reflect such changes without a major re-analysis of the system being analyzed. Proposed changes which are costly to implement or that may involve potentially dangerous situations for maintenance or operations personnel may be evaluated to determine their effective benefits prior to implementation of these proposed changes. The capability of sensitivity analysis also can provide the tools for optimization studies.

7. The technique should make use of a wide range of different data sources. There exists today, various types of data sources for human behavioral information. Data banks, field data, simulator data, laboratory studies and empirical and theoretical models of human behavior all have the potential for supplying needed data for human reliability studies. The development of any human reliability technique should not tie itself to any one source, but should be flexible enough to be able to use "good hard" data when they exist and to use the best empirical and theoretical models available when they do not exist.

8. The technique should be usable by non-specialists. The user of a human reliability prediction methodology should not be required to have expertise in the areas of human factors, psychology, sociology, equipment reliability, etc., in order to generate meaningful information from the model. Although these types of specialties may be required for the final analysis of a man-machine system, a human reliability technique should not require in-depth knowledge of their areas as a prerequisite for running the model. This factor not only aids in providing consistent results, but also eliminates some of the subjectivity involved in the application of the model by different users.

3.8. Quantitative Techniques for Human Reliability Assessment

Quantitative techniques for human reliability assessment have been primarily implemented in the operational context of system operation. Since little to no methodological development has been done specifically for the maintenance context, the currently existing *operationally* based methodologies were examined to determine their applicability to the maintenance area.

An in-depth review of the existing literature concerning human behavioral techniques has identified two general classes of predictive human reliability methodologies that may have some application in the maintenance area. These are analytic techniques and simulation techniques. Although other classes of reliability methodologies exist, i.e., mathematical modeling, the methodologies within the analytic and simulation techniques represent those methodologies which are most complete and most clearly defined. Because of this, the primary focus of this review will be on methodologies from these two classes of reliability prediction techniques.

Both analytic and simulation techniques employ some form of task/function/job analysis in order to determine the subtask/task units to be analyzed via the predictive technique. Each of the units is then analyzed to determine significant parameters that may influence the successful completion of the task/subtask. Data sources for the identified parameters are then compiled and utilized within the predictive technique to produce the human reliability metrics of interest. Although similar in rudimentary aspects, the two classes of techniques differ significantly in their approach of quantifying system failure rates resulting from human errors.

3.9. Analytic Techniques

Analytic techniques employ point estimates of the probabilities of success of the various task units and combine them mathematically according to some predefined scheme to derive point estimates of the reliability parameters of interest.

3.9.1. AIR Data Store

One of the earliest analytic methods developed for the prediction and evaluation of task reliabilities is the American Institute for Research's Data Store (AIR).[23,24] The data store contains information about various commonly found controls and displays for each item within the data store. Three types of information are listed for each item:

1. The probability of successfully operating the item as a function of its characteristics.

2. The minimum time required for operation of the item.
3. The additional increments of time needed to operate the item as a function of its characteristics.

The information within the data store was derived out of a review of over two-thousand research reports.

Use of the data store is relatively straightforward. A task analysis is done to determine the hardware/components that are involved in a performance of a particular task and the physical characteristics of these components. The appropriate human reliabilities (completion probabilities) with respect to the applicable component characteristics, are extracted from the data store and multiplied together to give the overall human reliability for the task being analyzed.

The data store method assumes complete independence of the probabilities used in calculating the overall reliability of the task. As a result, each multiplication of reliabilities tends to reduce the overall reliability. The method cannot account for feedback mechanisms within the task which can correct errors and increase overall task reliability. In addition, the probabilities used within the data store do not reflect the effects of the human characteristics of the operator. Characteristics such as training, prior experience, stress, fatigue, etc., may have a significant effect on the given probabilities. Environmental factors such as temperature and level of illumination may also affect the given probabilities and are not accounted for by this method.

The AIR data store was instrumental in establishing the basis of a human factors data base. Even though the method did not include various intrinsic aspects of tasks involving the human element, it did stimulate further research in the modeling of human behavior.

3.9.2. THERP

Probably the best known and most widely applied human reliability methodology is the Technique for Human Error Rate Prediction (THERP). Much literature by Swain and his co-workers have been written about this technique,[25,26,27,28] and its application.[29,30] Practical application has been done within the nuclear industry,[31] and a handbook[5] for applying the THERP technique to NPPs has been produced.

The THERP philosophy is similar to the philosophy used to predict overall system reliability via the AIR data store. Probabilities of various operations which make up a given task are multiplied together to produce overall system reliability. The THERP philosophy, however, goes beyond that of the AIR data store method, in that dependencies among various operations are accounted for by the use of conditional rather than independent probabilities. The THERP philosophy also allows human error probabilities (HEPs) to be modified by various performance shaping factors (PSFs). These PSFs allow special characteristics pertaining to the operator, his environment, and the operator's situation to be accounted for within the calculation of the overall system reliability. A list of various PSFs are given in Table 3.2. The PSFs in this table illustrate the large number of factors

Table 3.2. Performance Shaping Factors

External		Stressors	Internal
Situational Characteristics	Task and Equipment Characteristics	Psychological Stressors	Organismic Factors
Architectural features	Perceptual requirements	Suddenness of onset	Previous training/experience
Quality of environment: temperature, humidity, and air quality	Motor requirements (speed, strength, precision)	Duration of stress	State of current practice or or skill
Lighting	Control-display relationships	Task speed	Personality and intelligence variables
Noise and vibration	Anticipatory requirements	High jeopardy risk	Motivation and attitudes
Degree of general cleanliness	Interpretation	Threats (of failure, loss of job)	Knowledge of required performance standards
Work hours/work breaks	Decision-making	Monotonous, degrading, or meaningless work	Physical condition
Availability/adequacy of special equipment, tools, and supplies	Complexity (information load)	Long, uneventful vigilance periods	Attitudes based on influence of family and other outside persons and agencies
Manning parameters	Narrowness of task	Conflicts of motives about job performance	Group identifications
Organizational structure (e.g., authority, responsibility, communication channels)	Frequency and repetitiveness	Reinforcement absent or negative	
Actions by supervisors, co-workers, union representatives, and regulatory personnel	Task criticality	Sensory deprivation	
Rewards, recognition, benefits	Long- and short-term memory	Distractions (noise, glare, movement, flicker, color)	
Job and Task Instructions	Calculational requirements	Inconsistent cueing	
Procedure required (written or not written)	Feedback (knowledge of results	**Physiological Stressors**	
Written or oral communications	Continuity (discrete vs continuous)	Duration of stress	
Cautions and warnings	Team structure	Fatigue	
Work methods	Man-machine interface factors:	Pain or discomfort	
Plant policies (shop practices)	Design of prime equipment, test equipment, manufacturing equipment, job aids, tools, fixtures	Hunger or thirst	
		Temperature extremes	
		Radiation	
		G-force extremes	
		Atmospheric pressure extremes	
		Oxygen insufficiency	
		Vibration	
		Movement constriction	
		Lack of physical exercise	

that may influence the HEPs used within THERP. In addition to PSFs, THERP also accounts for the ability of an operator to make corrections to human errors that do not have an immediate effect on the operation of a system. These recovery factors (RFs) are used to modify the HEPs used within the technique.

Stress is one of the most important PSFs used within THERP. As can be seen from Table 3.2, Swain has utilized various types of stressors which may influence the level of stress experienced by the operator. THERP employs four levels of stress: relatively low, optimum, moderately high, and extremely high. The HEPs are directly affected by the level of stress experienced by the operator. Examples of its influence is evident in THERP's "halfing rule" which states that under optimum stress an operator takes extra care once he has made a mistake, and his HEP for the next step is one-half of his normal HEP. Under high stress situations, the normal HEP for the succeeding action is doubled when a preceding attempt has failed or when the action did not have its intended corrective effect. The effects of other PSFs on the HEPs are generally handled in a similar manner. In many cases, insufficient data exist to support the assignment of a particular value of a PSF and its effect on the HEPs, however, ranges of values, based on good subjective and well-founded common sense assumptions are used to substitute for missing data.

Application of the THERP technique consists of the following ten steps:

1. Describe the system goals and functions of interest.
2. Describe the situational characteristics.
3. Describe the characteristics of the personnel.
4. Describe the jobs and tasks performed by the personnel.
5. Analyze the jobs and tasks to identify error-likely situations and other problems.
6. Estimate the likelihood of each potential error.
7. Estimate the likelihood that each error will be undetected (or uncorrected).
8. Estimate the consequences of each undetected (or uncorrected) error.
9. Suggest changes to the system.
10. Evaluate the suggested changes.

A detailed discussion of the requirements of steps 1 through 10 is given in Swain's handbook[5] and are not repeated here due to space limitations.

THERP, like the AIR data store method, uses task analysis to describe jobs and tasks performed by the personnel (steps 4 and 5). The individual tasks, or steps in the tasks, become the limbs in the probability tree diagrams used for the reliability analysis. The analyst must use the information from the task analysis to determine the important PSF in deriving the probability estimates for each task or subtask.

The human error probabilities used within THERP are primarily what Swain calls "derived data." They may be extrapolations from related performance measures or a "best" judgement based on experience in complex systems or experimental and engineering psychology. Other data sources for input may be relevant reports and field calibrated simulator data.

Two primary measures are used by THERP.

1. The probability that an operation will lead to an error class i (P_i).
2. The probability that an error or class of errors will result in system failures (F_i).

The joint probability that an error will occur in operation and that it will lead to system failure is $F_i P_i$. If n_i, independent operations, are experienced, the probability of a failure condition existing as a result of class i errors is Q_i where:

$$Q_i = 1 - (1 - F_i P_i)^{n_i} \quad .$$

The total system or subsystem failure probability resulting from human error over all L classes of error is Q_t, where:

$$Q_t = 1 - \left[\prod_{k=1}^{L} (1 - Q_k) \right] \quad .$$

Q_t is interpreted as the probability that one or more failure conditions will result from errors in at last one of L classes of error.

The HEPs used within THERP to calculate overall system reliabilities will have been modified by various PSFs. One of the major criticisms of the THERP technique is with respect to the application of PSFs to human error probabilities. Successful application of PSFs requires a significant amount of expert judgement, familiarity with the necessary analysis techniques and a working knowledge of the system being analyzed. Unless the individual applying the THERP technique has had special training in its application, the use of THERP seems to be limited somewhat to highly skilled systems analysts. Another criticism of the THERP technique is that much of the data used is highly subjective in nature. When data or information is not available, extrapolations from existing data are made. Although these subjective extrapolations are based on assumptions that generally make good common sense, this information remains fragmentary and subject to biases in its application. Some individuals may view the heuristic nature of the technique coupled with the subjectivity of the data and the fact that documented validation of the technique does not exist (primarily due to the security classification of much of Swain's work), as drawbacks. The fact remains that THERP has gone far beyond the AIR data store technique in modeling human behavior and has effectively utilized existing data and information sources to a great degree in order to provide predictive reliability estimates.

3.9.3. TEPPS

Another analytic technique that is similar in some respects to THERP is the Technique for Establishing Personnel Performance Standards (TEPPS).[32,33,34] The technique developed by R. E. Blanchard and co-workers has two distinct modes — the derivative mode and the integrative mode. The derivative mode allocates system effectiveness requirements, such as the probability of system success, among personnel-equipment functional (PEF) units that are derived from a formal task analysis of the system of interest. The technique allocates the performance standards that are necessary to achieve the stated system goals. The derived standards can then be used as a measure of how effective actual performance standards are at meeting system goals.

The integrative mode is essentially a reliability model and is thus more applicable to the present literature survey. For this reason, discussions concerning TEPPS will be focused on the integrative mode.

As with the other two analytic methodologies discussed, a task analysis is performed on the system of interest to identify the PEF units which make up the system. Next, a function flow diagram called a Graphic State Sequence Model (GSSM) is produced. The GSSM is similar to THERP's probability tree diagrams and incorporates interdependencies among the PEFs and accounts for redundancies among personnel. Having completed the GSSM, the probability of successfully completing each of the PEF units is assigned. The probability data called the Index of Task Accomplishment (IOTA) is primarily from subjective sources and will be discussed later in this report. Once the IOTA have been assigned the GSSM is transformed via a computer program into a Mathematical State Sequence Model (MSSM). The MSSM is then solved by the use of the well known product rule to produce two primary outputs. These are: probability of task accomplishment and performance completion time.

The TEPPS methodology assumes optimal conditions for the system being analyzed. Because of this assumption, modifications to task performance and HEPs, such as THERP's PSFs, are not required and are not used. In addition no time constraints for successful task accomplishment are imposed. The probability of successful task accomplishment, however, is assumed to be a log-normal function of time between two time thresholds. These thresholds are T_{min}, below which there is a zero probability of successful task accomplishment, and T_{max}, above which the probability of successful task accomplishment is assumed constant. Other lognormality assumptions also exist within TEPPS. These include the number of individuals who perform a task successfully with respect to time and the probability of successfully completing combinations of tasks with respect to time.

Because of the lack of an abundance of human reliability data, the developers of TEPPS have made extensive use of subjective judgements to derive the required input data for the technique. Judges were provided with individual, detailed PEF descriptions and asked to compare each of the descriptions with all others in order to determine a ranking of the probability of successful accomplishment. This paired-comparison ranking technique provides meaningful distances or intervals between scale values and are used as the IOTA.

The scale values are transformed into probability values with the assumption that probability is an exponential function of the scale values obtained from the paired comparison technique. The exponential assumption used here by the developers of TEPPS was based on a regression analysis of probabilities within the AIR data store done by Irwin *et al.*, and reported by A. B. Pontecorvo.[35] The transformation requires an absolute determination of the probability of successful task accomplishment of the highest and lowest ranking scale valued PEF. From these values, two simultaneous equations are solved to provide the constants of the transform, which thus allows all other scale values to be transformed into probability values.

The primary criticisms of TEPPS centers around the exponential transform of scale values to probability values, and the extensive use of primarily one data source; subjective judgements. Both THERP and TEPPS make expert attempts at providing input data from subjective judgements in the absence of real data. Unfortunately, widespread acceptance of techniques utilizing subjective data probably will not come about until results can be validated via real empirical data. There has been only one reported validation attempt for TEPPS which proved unsuccessful.[34]

3.10. Simulation Techniques

The application of simulation techniques in the scientific, technological and military areas is not a recent occurrence. Since the advent of high-speed digital computers, various types of simulation models have been developed for a myriad of different uses. Siegel and Wolf[36] list 36 various simulation models that have been developed in various disciplines and represent a large spectrum of various simulation categories.

In the area of human reliability, simulation techniques have not been so readily applied. From a historical perspective, human reliability has primarily been approached with the same analytic techniques that were utilized in calculating equipment reliability. Simulation modelers, however, are strongly convinced that human behavior in a dynamic environment cannot be validly represented by deterministic methods. This difference in viewpoint coupled with the wide spread employment of traditional analytic approaches to human reliability has led to rather slow acceptance of simulation techniques.

Behavioral simulation models attempt to represent logically and mathematically, some category, type, or class of human behavior. They are concerned with person-machine, person-person, and person-environment relationships and with their interactions, i.e., with the performance of individuals and groups under varying conditions due to the environment, training, operational doctrine, characteristics of equipment systems and individual capabilities. In short, a behavioral simulation model is a dynamic representation of behavior and behavioral influences implemented on a digital computer so as to allow control or prediction of an event, or set of events.

3.10.1. Siegel-Wolf Models

The area of simulation modeling within the human factors field has been very heavily dominated by the work of Siegel and his co-workers. Siegel's work, which spans about three decades in this area, has been extensively documented with numerous reports and presentations (see Ref. 37 for a list of 25 reports and 13 presentations on digital simulation). In addition, Siegel and Wolf have published a book on man-machine simulation models.[38]

The primary digital simulation models of current interest come from what Siegel refers to as the family of task-oriented models. The three models to be considered from this family were developed to simulate individuals operating and/or maintaining equipment. In addition to major simulation variables, which reflect such items as the reliabilities of the equipment and mission reliability, the models include psychological and social variables (e.g., stress, orientation, proficiency, mental load, fatigue, etc.) pertaining to the operator or set of operators.

There are three primary models within Siegel's family of task-oriented models:

1. The 1-2 man model.
2. The intermediate crew size model.
3. The large group model.

Each of these models will be discussed later in this report.

As with all of the analytic techniques discussed, the task-oriented models rely on the performance of a task analysis to identify the subtasks involved within the operations or maintenance context. Once the subtasks have been identified, input data forms are completed which include information concerning task essentiality, precedence, average time for completion, standard deviation of the average time for completion, success probability and various personnel parameters such as stress threshold, speed/proficiency factors, etc. The larger-crew-size models include a greater number and variety of personnel characteristics than the smaller models. After specification of the input variables, the data are then translated into punched cards and entered into the computer. Each subtask is then sequentially simulated by the logic of the model and can account for operators working independently or together, waiting time, discussions among crew members, monitoring and operating controls and displays, skipping non-essential subtasks if the operators are busy, making decisions which can alter the subtask sequence, recycling (if required) in the event of an operator failure, operator incapacitation (partially or completely) and responding to unexpected failures and emergencies.

Simulation of the task is completed when the operator successfully completes the task within the allotted time or when all of the allotted time has been used. Each simulation of a task is called an iteration. Due to the stochastic nature of the simulation models, statistically meaningful results require a number of iterations of a particular task. Results from these iterations are then averaged to produce the final run results.

The computer output tabulations for the system of interest provides predictive data on system performance, personnel overloads, periods of unusual stress and excessive delays, distribution of how mission time is spent, a variety of end-of-mission conditions, and implications of manning strategies.

3.10.2. The 1-2 Man Model

The 1-2 man model[38] simulates one or two operators involved in a mission that may be minutes to hours long, involving up to 300 tasks for each operator. Each of the tasks may have a duration range of seconds or minutes. For each technician, four fundamental parameters are required:

1. *Stress Threshold* — Based on the concept that time-stress organizes behavior up to a threshold point and disorganizes it beyond that point, time-stress is defined to be the ratio of the average time to complete the remaining essential subtasks to the total time remaining to the technician. The effect of stress as a performance shaping factor increases the probability of successfully completing a subtask up to the stress threshold point. Beyond that point, the probability decreases.

2. *Time Available* — The total time available is a time limit for each operator for each task. All operators must successfully complete all of the essential task elements within their allotted time in order to achieve successful task completion.

3. *Individuality Factor* — Allows the effects of technician speed or proficiency to be incorporated into the simulation.

4. *Waiting Period* — Due to the nature of some types of cyclic subtasks, a technician may have to remain idle for a period of time which allows an initiated subtask to go to completion (e.g., allowing a chemical reaction to go to its completion).

A large amount of output data is generated for each iteration and for the entire computer run. The major output parameters are:

1. Task success probability.
2. Stress profile for the duration of the task for each technician.
3. Work, idle and failure-time summaries for each technician for the duration of the task.
4. Task repetition time.
5. The number of subtasks failed and ignored.

In addition to the task analysis, other data sources for Siegel's task oriented models are formal experiments, simulator measurements, literature reviews, personal interviews, and in general, the best available empirical and theoretical functional relationships.

The 1-2 man model is currently still in active use by a variety of government and commercial users. It has been applied to a number of different types of missions for validation purposes, e.g., carrier landings, missile launchings, in-flight refueling, etc. In all cases, the model's prediction ability compared favorably with independent data sources. A version of the 1-2 man model was also developed to simulate the electronics maintenance of Navy sonar systems. Siegel, *et al.*[39] reported that predictions made by this version of the model matched well with the independently obtained, empirically-based data for this situation.

3.10.3. The Intermediate Crew Size Model

The intermediate crew size model (ICSM)[39,40,41] simulates a crew of 4 to 20 men involved in a mission that may last for a duration of several days. A maximum of 80 tasks per technician per day can be simulated, and each of the tasks may have a maximum duration range measured in tenths of hours.

The ICSM includes many more parameters than the smaller 1-2 man model. Because of the increased crew size capabilities and increased list of applicable parameters, the ICSM is heavily group oriented. Information supplied to the model includes 88 input parameters concerning personnel, equipment repair, event and emergency event occurrences. A list of these parameters, their range, average valve and remarks concerning the parameters are presented in Table 3.3 and in Table 4-5 of the Human Reliability Prediction Systems Users Manual.[42]

The ICSM generates a crew of technicians or operators from personnel data that are entered. Parameters not entered are assumed to take on an average value. The model then assigns a value for speed, aspiration and competence to each crew member. The model also generates a daily schedule which includes scheduled, unscheduled and emergency events based on the input data. The model next selects an individual man or group of men to accomplish the work of each event. Depending on various constraints including time and event essentiality, non-essential tasks may be performed or ignored. Once the crew has been formed, daily schedules generated and assignment of personnel has taken place, the model proceeds to simulate each event. The model accounts for changes in psychological, physical and psycho-social variables due to the performance of each task. These and other parameters are evaluated and the next course of action is determined.

A large amount of output is generated by the ICSM. This includes printing the value of key variables, and summarizing end conditions for each event, each day, each mission iteration and a summary of all iterations. The major output parameters are:

1. Task success probability.
2. Stress profile.
3. Work, repair, idle and failure-time summaries.
4. Mean-time between failures.
5. Mean-time to repair.
6. Availability.
7. A determination of performance adequacy.
8. The number of tasks failed and ignored.

The ICSM has been in use since its initial development and validation in 1969. Since that time, the model has been improved to simulate situations oriented toward equipment, human and system reliability. Specifically, metrics of interest include human reliability,

Table 3.3. Intermediate Crew Size Model Input Data

Parameter Description	Range	Average	Remarks
1. Average aspiration	.5-1.0	.9	
2. Average crew pace	.5-1.5		
3. Average duration of scheduled event	.001-24		hours
4. Average psychological stress threshold	2.0-2.6	2.3	
5. Average repair time	0-24		hours
6. Average standard deviation of repair	0-999		
7. Average standard deviation of emergency	0-24		hours
8. Effectivity of stress	.5-1.0	.9	
9. Number of calories required by average crewman per day	2000-3500	2750	
10. Catnap length	.1-2	.3	
11. Duration time of emergencies	0-24		hours
12. Duration time of repairs	0-24		hours
13. Data change value	0-10		
14. Emergency event data set			
15. Repair event data set			
16. Class	1-20		
17. Description array			
18. Number of duty shifts			
19. Expected energy consumption	0-1000	250	calories per hr.
20. Expected energy consumption for emergency	100-1000	250	calories per hr.
21. Data change variable	0-10		
22. Family number	1-20		
23. Essentially	10-100		
24. Emergency essentially	0-100		
25. Essentially threshold	0-99	70	
26. Event type number	0-20		
27. Event number in family	0-200		
28. Event hazard class	1-9		
29. Event hazard class — emergency	1-5		
30. Printout option indicator array	0 or 1	1	
31. Event code	1 or 2		
32. Prerequisite event	0-200		
33. Equipment list	1-10		
34. Consumable rate of expenditure	0-9999		units/hr
35. Consumable rate of expenditure-emergencies	0-9999		units/hr
36. Number of repair events	0-200		
37. Physical capacitation fraction	0-1	.75	
38. Derating constant	.5-10	.9	
39. Event end type	1 or 2		
40. Initial level of rate of consumables			
41. Threshold consumable rate	0-9999		
42. Threshold consumables	0-9999		

Table 3.3. Continued

Parameter Description	Range	Average	Remarks
43. Mental load	1-9		
44. Mental load for emergency	1-9		
45. Maximum sleep	4-12	8	
46. Crew composition array	0-10		
47. Average number of man days per incidence of physical incapacitation	0-120	30	
48. Number of iterations	1-999	20	
49. Number of days	1-100		
50. Number of days between emergencies	0-1000		
51. Maximum number of days	1-30	5	
52. Duty shifts			
53. Number of family	1-10		
54. Equipment used array	1-10		
55. Number of scheduled events	1-200		
56. Number of men required by type	0-10		
57. Number of men required by type for emergency			
58. Next event number of each alternative	1-200		
59. Average duration of physical incapacity	0-30	3	days
60. Percent fully qualified in primary speciality	0-100	25	
61. Percent moderately qualified in primary speciality	0-100	50	
62. Percent unqualified in primary speciality	0-100	25	
63. Probability at each alternative path	0-1.0		
64. Cross training probability table	0-1.0		
65. Average short term power output	200-500	350	calories
66. Equipment reliability	0-1000		days
67. Intermittent reliability	0-1.0		
68. Repair touchup code	1-3		
69. Sea state	0-9	3	
70. Standard deviation of body weight	0-50	15	pounds
71. Number of hours since last eight hour sleep period	0-20	7	
72. Percent fully qualified in secondary speciality	0-100	25	
73. Percent minimally qualified in secondary speciality	0-100	50	
74. Percent unqualified in secondary speciality	0-100	25	
75. Earliest starting time allowed	0-24		hours
76. Fatigue threshold	1-.35	.25	

Table 3.3. Continued

Parameter Description	Range	Average	Remarks
77. Time limit by which event must be completed	0-24		hours
78. Consumable threshold set identifier	0-9999		
79. Consumable set identifier	0-9999		
80. Threshold set rate for consumables below which event is ignored	0-9999		
81. Threshold set for consumables below which event is ignored	0-9999		
82. Threshold set rate for consumables below which emergency is ignored	0-9999		
83. Threshold set for consumables below which emergencies are ignored	0-9999		
84. Intermittent reliability	0-1000		hours
85. Number of hours worked after which no new work assignment is made	4-20	16	
86. Number of hours worked after which no new work is authorized	4-16	10	
87. Mean body weight	1-300	150	pounds
88. Physical capability constant	0-100	25	

availability, and mean time to repair; equipment reliability, availability, mean time to repair and mean time between failures; and system reliability, availability, and mean time to repair.

Validation testing of the ICSM was carried out by simulation of a 4-day mission of an operational Viet Nam river patrol mission. The mission included 60 to 65 events in all, and the simulation results compared favorably with results obtained from interviews conducted with Coast Guard officers.

3.10.4. The Large Group Model

The large group model (LGM)[38,43] simulates a crew of 20 to 100 men involved in a mission that may last many days. A maximum of 100 tasks per technician per day can be simulated, and each of the tasks may have a duration range measured in hours.

The LGM is primarily concerned with group performance, and inputs to the model are principally concerned with group oriented variables. Crew size and composition may be pre-specified, or if left unspecified, the LGM will calculate a sufficient, minimum crew size for simulation. The LGM will automatically increase the crew size during each iteration by selecting and adding the most needed man or men to the previous iteration crew. Additions continue to be made until a maximum specified crew size is reached, or until the crew reaches a size which eliminates the need for overtime work.

The LGM requires an extensive amount of input data and generates output summaries comparable to the ICSM. The major output parameters are:

1. Task success probability.
2. Crew efficiency.
3. Crew morale.
4. Work, repair and failure time summaries.
5. The number of tasks failed and ignored.

Operational validity of the LGM was demonstrated by performing numerous simulations of a 21-day submarine mission with crew sizes varying from 48 to 61 men. The results compared favorably with data obtained from independent sources.

Of all the simulation models discussed thus far, the LGM is the only one that is currently not in active use. The reason for this is probably the large amount of input data that is required for the model to run.

3.10.5. SAINT

SAINT (Systems Analysis of Integrated Networks of Tasks)[44,45,46,47,48] is a Fortran-based, network-oriented, combined simulation language for modeling large, complex systems. Although SAINT is not a human behavioral methodology per se, it will be discussed briefly within the simulation section because of the degree of flexibility that is offered by this language to analysts wishing to employ simulation techniques.

The origins of SAINT can be traced back to two distinct technologies; task analysis and the Siegel-Wolf[49] simulation of operator performance under workload stress. These technologies provided the basis for identifying and incorporating many of the currently existing features in SAINT. Subsequent versions of SAINT adapted many features of GASP-IV (General Approach to Systems Programming),[50] another simulation language, and resulted in a flexible, sophisticated, combined modeling technique where networks of discrete events could be modeled along with the dynamics of continuous processes.

The SAINT language offers its users a method whereby systems may be modeled within a simulation methodology format without the requirement of an in-depth knowledge of simulation techniques. The simulation model developed through SAINT, however, requires that the developers have an in-depth working knowledge of the attributes, characteristics and parameters comprising the system being analyzed. This may require the resources of a team of individuals consisting of psychologists, systems analysts, human factors experts, sociologists, engineers, etc. in order to effectively model the system of interest. The actual expertise needed depends on the complexity of the man-machine system being modeled, and the depth to which the model is to be developed.

Because SAINT is a flexible modeling tool, applicable in theory to a large spectrum of various system contexts, it offers its users a wide variety of modeling options. One of the most flexible options in SAINT is the ability to represent the dynamics of the hardware of the system being analyzed. Once the dynamics of the hardware have been determined, they may be incorporated into SAINT via a moderator function. The moderator function is a Fortran subroutine written by the user representing the dynamics of the system under analysis. The use of moderator functions is not restricted to representing the dynamics of system hardware. It may also be used to dynamically alter any system variable as the code is running, as long as the proper functional relationship has been represented within the moderator function.

SAINT also offers its users a choice of four types of branching logics. These are:

1. *Deterministic*, whereby all branches emanating from a task are selected on task completion.
2. *Probabilistic*, whereby each branch emanating from a task has an associated probability of being selected, and only one branch is selected on task completion.
3. *Conditional — Take First*, whereby each branch is associated with a condition and are ordered. Each condition is tested sequentially in the specified order and the first branch with a satisfied condition is chosen.
4. *Conditional — Take All*, whereby each branch with a satisfied condition is chosen.

SAINT also offers to its users a wide variety of choices for specifying the time to perform a task. Eleven different distributions are available which are: beta fitted to three values, constant, Erlang, gamma, log-normal, normal, Poisson, triangular, uniform, and Weibull. Additionally, SAINT facilitates contingencies, decision making and emergency conditions via its attribute assignment or branching logic (mentioned earlier).

Discrete and continuous task interactions are handled via threshold values of the state parameters associated with a continuous task. Threshold crossings by state variables can signal or initiate tasks. They may thus influence task performance characteristics and precedence relationships. Threshold crossings can also change the values of logical variables which in turn may alter the form of equations or their precedence.

The most common use of SAINT is in the study of resources, performing tasks to achieve a desired objective. Because of this, SAINT automatically generates, for the user, utilization statistics on all defined resources. In addition, SAINT also allows the user to generate various types of statistics concerning the tasks to be accomplished. Four different statistic types, i.e., the time of the first occurrence of a task; the time of all occurrences of a task; the time between occurrences of a task; and the number of occurrences of a task may be collected at 4 types of occurrences, i.e., the occurrence of the release of a task (ready to be initiated), occurrence of the start of a task, occurrence of the completion of a task, the occurrence of the interruption of a task in progress. In addition to these 16 task specific statistics, SAINT also allows interval statistics to be generated which provide information concerning the interaction or relations between tasks. Interval statistics may be initiated at the release, start, completion or interruption of a task. Other user requested SAINT outputs include graphic portrayals of the probability and cumulative density functions for a distributed variable and a time trace of specified state variables.

Utilization of SAINT combines system information from a task analysis or other information sources with the experience and knowledge of a systems analyst. Through the use of a set of standardized SAINT modeling symbols, the systems analyst develops a graphic network model. The network model identifies each task to be performed, the resources required to perform the tasks, the state variables, the conditions that initiate interactions between state variables and task performance and the environmental conditions under which the system is required to operate. The modeling symbols used within the network model were designed to allow easy preparation of the data input cards required to run SAINT. Once these cards have been prepared and submitted to the computer, SAINT generates the desired measures of system performance.

As a modeling tool, SAINT offers the analyst a systematic approach for describing the man-machine system to be analyzed. It offers the capability of modeling a wide variety of complex systems with an extensive set of user specified options. Again, it must be emphasized that SAINT is not a model, but a language which provides the framework within which any proposed model may be described. Limitations of applying SAINT essentially lie in the capabilities of the modeler and his ability to quantitatively describe the system of interest. Human variables such as fatigue, stress, motivation, etc. have the capability of being brought into the simulation process via the moderator functions. These must be specified by the user and are only as good as the functional relationships that the user specifies. Thus, although SAINT allows simulation modeling to be used by individuals who are not experienced in simulation techniques, the effective application of SAINT to tasks involving a large number of characteristics that must be specified via a moderator function, requires a significant expertise in human factors, psychology, and other applicable areas. SAINT in effect, removes from the model developer the requirement to be expertly versed in simulation modeling, but does not remove the associated requirements of expert knowledge concerning the effect of various types of performance shaping factors.

3.11. Comparison and Summary Discussion of Methodologies

Experience with NPP operation has shown that human errors committed within the maintenance context of system operation can contribute significantly to the overall risk from power plant operation. Human factors work within the maintenance context to date has primarily taken the form of general human factors reviews in an effort to increase maintenance personnel effectiveness. Little emphasis has been placed on the quantitative assessment of human reliability within the maintenance context.

Human factors work within the operational context, however has resulted in the development of quantitative, predictive operator reliability techniques. Results of a literature review of these models included the identification of two general classes of predictive human reliability methodologies — analytic and simulation techniques. The analytic techniques examined included the AIR data store method, THERP, and TEPPS. The simulation techniques included the three Siegel-Wolf models and the SAINT simulation language. One of the primary differences between these two classes of techniques lies in the method with which each technique approaches the problem of task interdependencies. Analytic techniques employ the well-known product rule with probabilities that may have been modified by various performance shaping factors. Simulation techniques provide dynamic representation of the system of interest which evaluates the resources available to determine subsequent courses of action. Probabilities of successful task completion are modified internally depending on the state of the dynamic variables in the model.

Both analytic and simulation techniques require the resources of some type of data base. The lack of a sufficient base of objective empirical data in the operational context has necessitated the subjective extrapolation of existing data to meet the informational requirements for tasks of which no data exists. In addition, other "soft" data sources such as expert opinion are often used to fill the gap of unavailable data. Although it can be argued that "soft" data sources are no better than guesses and should not be utilized, the fact remains that at present these data sources represent the best that is currently available and are utilized effectively by both analytic and simulation techniques.

The form of the probability data used within the two classes of techniques differ significantly. Analytic techniques utilize point estimates of human error probabilities and associated performance shaping factors. Simulation techniques utilize probability distribution functions with distribution parameters based on the best existing data. Neither demonstrates a distinct advantage over the other.

Other methodological differences between the two classes of techniques include the mechanisms whereby task sequencing, task ommission, feedback and emergency scenarios are handled. Although one of the analytic techniques — THERP — has accounted for feedback by the use of a performance shaping factor (the recovery factor), simulation techniques have a distinct superiority in their ability to account for these mechanisms.

The requirements by each of the techniques on the user of the technique is also somewhat different. Analytic techniques (e.g. THERP) require that the performance shaping factors which influence the human error probabilities be specified by the user. This requirement necessitates that the user be familiar with how each of the important performance shaping factors vary as situational circumstances change. No such requirement exists for simulation techniques because interdependencies are handled dynamically by evaluating internally the important model variables. Simulation techniques are thus not only usable by experts in the human factors field, but also usable by non-specialists.

In order to quantitatively assess the contribution of maintenance activities to the overall risk associated with the operation of a nuclear power plant a thorough understanding of the many variables involved and their interactions must be obtained. A requirement for understanding the complex interactions of these variables is that the systems analyst has an underlying structure or model for their relationships. Currently, no comprehensive conceptual structure or model exists for maintenance activities, and the currently existing methodologies for the operations context of systems operations were examined to ascertain their applicability to the maintenance area.

Information obtained through the review of the various methodologies indicated that a potential for application to the maintenance area existed, however, as they currently stand, none were directly or generally applicable to maintenance tasks performed at NPPs. Maintenance tasks generally are either very labor intensive, or require a great deal of cognition (during troubleshooting), both being uncharacteristic of tasks in the operational context of systems operation, for which existing methodologies were developed.

Throughout the review of the existing methodologies, each technique was evaluated with respect to the eight specific needs of a quantitative, predictive technique presented earlier in this report. In addition, various features of each of the methodologies were examined for applicability in the maintenance area. The results of this evaluative process lead to the selection and recommendation for development of a simulation technique directed specifically toward maintenance activities in NPPs. Aspects of the evaluative process will now be presented.

Both analytic and simulation techniques have the capability of handling various types of operator positions and a mixture of positions in a crew structure. Interdependencies between task elements are handled much more effectively by simulation techniques than analytic techniques. Simulation techniques handle interdependencies between task elements by means of dynamically evaluating various important model parameters and determining internally the effect on subtask success probabilities. Analytic methods, on the other hand, utilize conditional probabilities and combine the resultant probabilities via the well-known product rule. All of the techniques examined with the exception of the AIR data store accounted for both human and equipment characteristics in their respective methodologies. Simulation techniques provide a wide range of output information on both the subtask and task levels. Important parameters may be tracked throughout the simulation process and output distributions generated. Analytic techniques generate only one or two global parameters such as probability of successful completion of a task and overall system reliability. Simulation techniques thus can provide the data analyst with detailed information that lends itself easily toward understanding, on a broader basis, the complex interactions and dependencies which exist during maintenance activities.

The ability to model tasks which are highly cognitive in nature involves the consideration of many types of interdependencies. As was mentioned earlier, interdependencies are handled much more effectively with simulation techniques than analytic techniques. Although no existing human behavioral methodologies currently include a comprehensive technique for handling cognitive task elements, simulation techniques seem much more suited for their inclusion than analytic techniques.

A valuable tool that may be used to ascertain the impact of changes in system design, maintenance procedures, training, team composition and structure is sensitivity analysis. The analytic techniques examined in this review do not lend themselves easily to sensitivity analysis. The AIR data store assumes independent success probabilities, while TEPPS assumes the existence of optimal conditions for the system being analyzed; both unacceptable with respect to sensitivity analysis. THERP with its use of performance shaping factors does have the capability of doing sensitivity analysis. However, for each case to be run in a sensitivity analysis, changes to the performance shaping factors must be evaluated extrinsically and may be biased on a subjective basis by the systems analyst involved in ascertaining their change. Because simulation techniques perform an internal resource evaluation prior to the simulation of a task, no additional systems analysis needs to be done in the running of sensitivity analysis cases. The parameter or parameters to be varied are simply changed to their new value in the input data. Simulation techniques are thus much more amenable to sensitivity analysis than analytic techniques.

The ability to utilize various types of data sources as input to a human behavioral model is a valuable and necessary asset, especially when an abundance of real data does not exist. The developers of TEPPs were aware of the lack of human reliability data and made extensive use of subjective judgements to derive the required input data for the technique. Data sources for TEPPs were limited almost exclusively to this type of information. In THERP, human error probabilities and performance shaping factors are both susceptible to change when new field data or information from new empirical or theoretical models on human behavior are developed. Simulation methods are also affected in a similar way. Both Swain and Siegel advocate the use of the best data available in running their respective model, and both have emphasized that if a user has better data available, it should be used.

Any human reliability model that is to enjoy widespread application and use must be usable by non-specialists. The requirement of highly specialized prerequisites for running the model may lead to limited use and perhaps misuse. Although in-depth analysis of human behavior requires highly specialized, trained individuals, the use of a model should not require that the user be expert in *all* associated discipline. The degree of speciality required by the various techniques that were examined varied considerably. The AIR data store is conceptually simple and requires a minimum of specialized knowledge in determining the effect to success probabilities due to varying equipment characteristics. Because TEPPs operate in an optimal mode, it also requires very little specialized knowledge. The generation of the subjective success probability data required by TEPPs does necessitate becoming familiar with mathematical transforms and paired comparison techniques. Nevertheless, an abundant amount of specialized knowledge is not required to run TEPPs. THERP makes extensive use of performance shaping factors which modify the human error probabilities used in the model. The assignment of performance shaping factors by the user requires a significant amount of expert judgement, familiarity with the necessary analysis techniques, and a working knowledge of the system being analyzed. The user must have a good feeling for how the various performance shaping factors are affected by changes that may occur in the man-machine system. The use of THERP, thus seems not to be geared toward use by non-specialists. Simulation models, especially the larger crew sized models, required a large amount of input data. Although the compilation and input of this data into the model requires a good deal of effort, no specialized expertise is needed in running these models.

The literature review of human behavioral methodologies did not reveal any documented comparisons between the techniques discussed in this report. The choice of the methodology to be developed for the nuclear power plant maintenance context was thus made on the basis of how well each of the methodologies fared with respect to various identified specific needs, and unique characteristics offered by the difference methodologies. Simulation methodologies have the capability of meeting many of these needs. In addition, the methodologies in the simulation class have been the only model that have undergone validation studies with positive results. Simulation methods also offer unique methods of altering the sequence of task performance, incorporating emergency situations among the maintenance scenarios, the ommission of non-essential tasks and the ability to account for the potential of error correction by the human.

On the general level, the advantages of computer simulation over more deterministic analytic techniques appear to be that simulation allows the capability for:

1. Analysis and prediction which considers the inherent variability both between and within individuals. This feature seems particularly important to nuclear power plant analysis where requirements and proficiency may vary both across and within plants.
2. Considering a large number of equipment, environmental, and personal variables independently and is continuous. The number of possible interaction in large complex systems such as a nuclear power plant is very large and cannot be handled efficiently by other analytic methods. In addition, trade-off curves between variables can be generated.
3. Providing prescriptive (what should be done) as well as diagnostic (what seems to be the problem) information.
4. Considering in a realistic manner both the "normal" and the "degraded" condition for both the personnel and the equipment in a system.
5. Provides results in the form of distribution rather than as point estimates.
6. Considering a current system, hypothetical variations of the system and alternative systems.
7. Allowing for random events to be superimposed on event sequences.
8. Provides results of a variety of types and at a variety of levels. For example, completion time, success probability, error rate, and error type data can be provided by person, shift, day and extended time period.
9. Dynamically considering the time varying characteristics of humans. For example, humans learn, become fatigued, gain or lose motivation, and respond to stress.

Because of the desirable model characteristics identified during the front end survey and the desirable traits of simulation modeling identified through the literature review, it is recommended that if a formalized, quantitative, predictive methodology for the analysis of nuclear power plant maintenance personnel reliability is to be developed, it should be of the simulation type.

4. FUNCTIONAL REQUIREMENTS OF A MAINTENANCE MODEL

4.1. Introduction

The results of the user interviews mail survey and literature review described in the previous chapters supported the conclusion that the important aspects of the nuclear power plant maintenance activity should, in fact, be identified and incorporated into a simulation model. Certainly, the respondents of the mail survey concurred in the general utility of the information which could be provided by such a model.

This chapter presents summary or functional requirement statements for such a simulation model. These requirements, while categorized by primary subject matter area, are not ordered in terms of priority. They serve as top level guidelines from which more detailed requirements and designs can evolve for a model for simulation of the behaviors involved in nuclear power plant maintenance. Some items are stated as desirable to indicate a lesser requirement than mandatory. Generally, the substance of the requirements are based on the results of the mail survey. Additionally, computer programming and functional user-oriented requirements are presented.

4.2. General Requirements

- The model will be developed in the form of flow logic suitable for programming on a general purpose digital computer.
- The model will be a behavioral simulator suitable for the prediction of nuclear power plant maintenance results for one maintenance task at a time.
- The model will be applicable to any corrective as well as any preventive maintenance task performed at any nuclear power plant with the following limitations:

    a. task duration is no longer than two days

    b. the maximum size of the maintenance team at any one time is eight

- The model will be applicable to maintenance performed by nuclear power plant personnel and (desirable only) to contractor personnel.
- The model will be stochastic in that it will incorporate Monte Carlo features to simulate chance and choice elements.
- The model will be capable of supporting sensitivity testing and validation testing.
- The model will permit simulation of up to five types of maintenance personnel including:

    1. mechanical
    2. electrical

3. instrument and control
4. supervisor

- The model will be utilized by selecting run conditions and initiating a computer run (for a user-selectable number of simulation iterations) of the selected task and monitoring printed results, as well as output displayed on a cathode ray tube terminal.
- The model will simulate the replacement (rotation) of maintenance personnel during task performance.
- The model will simulate maintenance tasks for existing as well as planned nuclear power plant facilities, i.e., it will support nuclear power plant developmental efforts from the preliminary design through implementation phases.
- The model will simulate (up to 100) subtasks comprising a simulated maintenance task on each task simulation iteration.
- The model will simulate on each subtask the assignment of a group of maintenance personnel composing a team consisting of one to all eight persons.
- The model will simulate person-to-person as well as person-to-machine oriented subtasks.

4.3. User Oriented Requirements

- The model will incorporate interactive conversational features to facilitate ease of use by the user.
- The model will be operable from an alphanumeric computer terminal.
  a. the terminal shall have a capacity of at least 24 lines of 80 characters (desired 132 characters)
  b. the terminal will provide a direct hard copy record of any CRT display
  c. it shall be possible for the user to control input, modify any display, and/or print any user input data from the terminal
  d. the terminal will support limited graphic displays (desirable).
- The model will accept input from the user entered via menu responses.
- The model will provide a response to the user (either the requested result/display, or a "request in process indicator" within three seconds, 90 percent of the time.
- Wherever possible, the model will provide default data when inputs are not provided by the user, and accept team oriented inputs instead of individual inputs at user option.
- The model will provide outputs in organized printed form from a computer line printer.
- The model will display error indicators to the user on the terminal (i.e., input sequencing, value outside permissible range, etc.).
- The program will provide assistance to the user by display of:

a. menus for user-initiated options/actions

b. a list of tasks for which input data are available.

- Selection of data element names will be based on ease of recall.

4.4. Computer and Programming

- The computer program will be capable of meeting all functional model requirements specified.
- The computer selected shall support all of the following:

a. the FORTRAN IV or a simulation language (GASP, GPSS, SIMSRIPT)

b. pseudo random number generation

c. terminal and peripheral requirements (see user requirements)

d. terminal log-on and security features

e. program, tracing, and debugging features.

- The program will be organized into interrelated subroutines or modules, each of which possesses logical and processing coherency.
- The program will be developed on a top down basis using structured programming techniques.
- Selection of data element names will be based on ease of recall.

4.5. Input Requirements

- The model will permit simulation of maintenance using any combination of input variations itemized below.
- The model will allow for full simulation of within and between individual differences.
- The model will operate using a realistic set of user-selectable parameter inputs for features including the following maintenance conditions:

a. the environment including radiation levels, and temperature (desirable only: acoustic noise level)

b. maintenance team work orientation (work attitude)

c. human factors situation, i.e., tight places

d. task time limit (time constraints)
e. worker's aspiration
f. maintenance team ability levels
g. team proficiency (prior experience)
h. team training and composition
i. safety gear.

- The model will utilize task analysis data for each simulated task consisting of information in the following categories for each of up to 100 subtasks comprising the maintenance task:

a. subtask essentiality
b. next subtask identifiers
c. specific types of personnel assigned
d. kind of subtask
e. protective clothing status, i.e., safety gear
f. waiting and idling requirements (precedence indicators)
g. radiation absorbed dose
h. fatigue, stress
i. time worked, idle, and waiting
j. safety level
k. task success probability
l. training.

- Other input data will include:

a. task identification
b. run identification selected by the user
c. limit values for selected variables beyond which special output flags are displayed
d. required ability levels for each subtask and/or kind of subtask. These will be employed to simulate and predict undetected errors.

4.6. Calculation Requirements

- The model will calculate average subtask and task success probabilities.

- The model will include calculation of psychological, physical, and social interactive variables upon which principal outputs will depend. These variables and their functional relationships with inputs, outputs, and each other shall be selected in accordance with current, accepted human factors and psychological theories and in accordance with results of the user survey.
- The psychological variables will include some aspect of at least the following elements for the maintenance personnel:

  a. stress

  b. fatigue

  c. maintenance personnel ability, including mental and mechanical ability levels

  d. worker attitude.

- the model will include all necessary logic and programmers for initialization, storage, selection, and update of all selected variables.
- The program will control user input of data, and respond to user requests including output listings and user displays.
- The program will provide for the possibility of repeating subtasks (looping) and of skipping subtasks.
- The program will detect and report logical inconsistencies in input data and other error conditions to the user.
- The program will allow variation of input conditions by the user one at a time or in any combination.
- The computer selected and its program will have as a goal the execution of a task simulation (consisting of 50 subtasks for 10 iterations) within two minutes.
- The model will meet the objectives stated here, but its design will be based upon reasonableness of output volume, acceptability of values over the simulation range, processing economy, indifference to trivial effects, and the relative importance of output results as described in the survey.

4.7. Output Requirements

- The model will generate a variety of outputs either on a line printer located at a (remote computer site) or on a local cathode ray tube screen and its hard copy device:

  a. detailed subtask level output (user option)

  b. end-of-iteration and end of shift results (user option)

  c. end-of-run summaries

  d. input data listings

- The programmed outputs will include summaries of run simulation results including end-of-iteration and end-of-run averages, peak or out-of-tolerance conditions, and frequency distributions. They will represent the results of the simulation under conditions of the selected inputs. Provision will be included for results displayed by subtask, task, or both.

4.8. Documentation Requirements

- The computer program listing shall be fully commented.
- All information required by a user for efficient running of the model will be presented in a step-by-step manner in a user's manual.
- A model report containing functional relationships, model flow diagram, data item name, program description, module descriptions, and input and output formats will be issued.

4.9. Training Requirements

- The model shall be sufficiently simple in organization so that:

    a. a demonstration of all of its user-oriented features from a terminal can be presented within three hours

    b. a course in learning to run the model from a terminal can be completed in two days (excluding subtask data preparation) by a person of Bachelor of Science (engineering, science) level.

4.10. Summary and Conclusions

Chapter 4 of this report has presented general functional requirements for a simulation model that addresses maintenance activities at nuclear power plants. The requirements presented were based generally upon the results of the front-end user survey and identified needs and characteristics that are highly compatable with the abilities of simulation modeling.

Chapter 3 of this report examined existing human behavioral methodologies for the purpose of selecting the type of methodology for this program which would be most suitable for the needs of its potential users. The methodologies were evaluated against a set of desirable model characteristics, and the choice of simulation modeling was made primarily because of its superior ability to handle interdependencies, provide point estimates and distributions of output variables, was capable of being computerized and has been validated for other contexts.

The choice of simulation modeling for this program seems to be supported by the results of the independently conducted front-end user survey. Many of the needs identified through the survey can best be handled by simulation modeling. In addition, simulation modeling will provide a systematic and logical means of addressing the maintenance context. The highly interdependent nature of many of the activities of maintenance personnel will be put into a simulation framework which will allow analyses that were previously difficult or impossible to be carried out. The dynamic nature of simulation modeling allows the many interdependencies to be handled in an efficient, effective, and realistic manner.

Because of the characteristics identified during the front-end user survey and because of simulation modeling's superior ability to address complex, interdependent human activities, it is recommended that development of a simulation model for the maintenance context should be persued. Chapter 5 of this report provides a comprehensive plan for the development, validation, and dissemination of such a model.

5. PLAN FOR NUCLEAR POWER PLANT MAINTENANCE MODELING

5.1. Introduction

Chapters 1 through 4 of this report have addressed the need for, the type of and the functional requirements of a methodology that can provide reliability measures of human performance within the maintenance context. The remaining sections of this report will provide details concerning efforts that will be required to develop, validate, and disseminate the proposed model. Appendix E of this report discusses in further detail a program plan to accomplish the efforts described in this chapter, including a detailed schedule of tasks and subtasks to be performed.

5.2. Users and Uses of Maintenance Performance Reliability Information

The Nuclear Regulatory Commission (NRC) has been charged by the federal government with the responsibility of assuring the safe and reliable operation of Nuclear Power Plants (NPPs) within the United States. One aspect of this responsibility is to ascertain and reduce the risks associated with the operation of these plants. In particular, NRC's programs within the Probabilistic Risk Assessment (PRA) area are endeavoring to identify, quantify, and provide guidance for minimizing any risks associated with a NPP.

Any comprehensive program dealing with the quantification of risk from a total systems perspective must include the human as an integral part of the system being analyzed. The inclusion of the human element within these analyses necessitates the existence of a comprehensive data source that can provide needed reliability metrics related to human performance.

It is precisely because of this need and because of the sparcity of comprehensive data sources related to NPP maintenance activities, that the NRC is sponsoring the development of a model that will provide estimates of the reliability of the performance of NPP maintenance personnel.

The proposed model is being developed primarily for use by the NRC, especially for support of the human reliability analysis portion of probabilistic risk assessment. Because of the model's intended generality, and because it will provide a logical and systematic means of analyzing maintenance tasks in terms of the numerous interdependent variables associated with these tasks, it will have potential benefit to a large number of users with different needs.

In addition to its use for PRA purposes, the model will be of use to NRC as an aid in prioritizing research within the maintenance area. That is, the model will be able to aid in the identification of existing and potential problem areas within maintenance and provide a quantitative basis for prioritizing these problem areas for future research. The model may also be utilized to supply information valuable to the selection and training of maintenance personnel as well as to provide input to the NRC in establishing standards for maintenance personnel qualification.

The nuclear utilities may also benefit from information provided by the proposed model. In particular, the model will be able to aid in providing information relevant for maintainability design and maintenance organizational structure at a nuclear facility, as well as aiding in decisions concerning staffing levels.

Information that will be provided by the proposed model may also be useful for the allocation of functions within the maintenance area as well as the evaluation and improvement of job performance aids. In general, the proposed model will provide benefits to any organization requiring information concerning the performance of maintenance activities at a Nuclear Power Plant. Some of the primary uses of the proposed model are listed in Table 5.1.

5.3. Description of the Model

This chapter presents the initial concept and structure of a model whose purpose is the prediction and evaluation of nuclear power plant maintenance performance reliability.

The model is named MAPPS, for *M*aintenance *P*ersonnel *P*erformance *S*imulation.

5.3.1. Model Overview

The global purpose of the MAPPS model is to allow quantitative analysis of the effects of varying a set of conditions represented by model inputs (called set A) on a second set of conditions or analytic results (called set B). Examples are shown in Table 5.2.

The MAPPS model possesses the capability for generating descriptive and predictive information which will be useful in a variety of ways to those who are concerned with nuclear power plant maintenance and particularly those interested in public safety and public risk analysis.

The input conditions can be varied one at a time or in any combination by the user at a computer terminal. The outputs are provided at various levels of detail as selected by the user. Generally, all of the outputs are available in summary form. In this way, a user can design his numerical experiments consisting of one or more runs* and be presented with data representing all elements of set B outputs from which he can develop relationships, gain interdependency insights, and draw hypotheses and conclusions about various aspects of nuclear power plant maintenance.

5.3.2. General Features of the Proposed Model

The top level features and requirements of the MAPPS simulation model were defined on the basis of a front-end analysis of user needs and requirements (Chapters 1 and 2 of this report) and the job analysis of two nuclear power plant maintenance positions.[51,52]

*A run is the summary of several "simulations" of the same initial conditions.

Table 5.1. Primary Uses of the Proposed Model

Reliability/Risk/Safety Assessment

- Estimate human reliability in existing maintenance activities
- Estimate human reliability in proposed maintenance activities prior to implementation
- Identify "error-likely situations"
- Estimate improvements in human reliability due to modifications in equipment, procedures, training, etc.
- Identify and estimate the effects of critical human variables, e.g., fatigue, stress, experience, etc.
- Estimation of accumulated personnel radiation dose

Maintenance System Design Evaluation

- Estimate the effectiveness (e.g., mean time to repair) of existing systems
- Identify potential maintainability problems in existing systems
- Integrate human engineering and personnel-equipment designs into maintenance system designs
- Identify personnel requirements (section and training)
- Develop and evaluate maintenance procedures

Maintenance Operations Analysis

- Comparison and optimization of maintenance strategies
- Optimization of team assignments (type, number)
- Maintenance planning/scheduling

Human Factors Data Store

- Contribute human factors data and information for maintenance to a human factors data store

A summary enumeration of the planned model features are as follows:

1. MAPPS will be a general model which is applicable to any nuclear power plant. It will have the capability of simulations:
   a. corrective as well as preventive maintenance tasks
   b. contractor as well as "in house" maintenance
   c. repair conducted by personnel of any combination of skills who are working under any conditions usually encountered

Table 5.2. Examples of Parameters That May Be Varied and of Analytic Results

Set A — Inputs	Set B — Results
Team composition	Task duration (MTTR)
Team proficiency	Where errors are made
Radiation exposure rate	Undetected errors
Temperature level	Risk weight
Use of safety/protective gear	Performance effectiveness
Light/crowding/accessibility level	Total radiation absorbed dose
Time constraint (limit)	End of task fatigue, stress
Team stress level	Team member time worked, idle
Team work orientation	Team safety level
Team training level	Task success probability
Time of day	Team reliability
Prior hours worked	Training effects
Team abilities	Skill retention
Stress threshold	
Task load	
Adequacy of procedures	

2. MAPPS will simulate any single maintenance task which:
   a. has a duration between 15 minutes and two days
   b. can be accomplished by a "team" of eight or fewer maintenance personnel

3. MAPPS will allow for:
   a. variation of technician ability level, level of training, environmental and situational variables, and the human factors situation
   b. customized task analysis data for each simulated repair task including such data for each subtask as essentiality, next subtask identifiers, men assigned, and subtask precedence conditions
   c. calculated (not input) values for average subtask durations and success probabilities
   d. a taxonomy of maintenance team ability level factors

4. MAPPS will include provisions for:
   a. using default data when selected inputs are not provided by the user
   b. using a Monte Carlo approach for several functions to introduce chance elements
   c. replacement (rotation) of maintenance personnel

d. operating via interactive computer terminal; providing hard copy results at the terminal and on a local computer line printer

e. calculating values of pertinent variables such as stress, fatigue, and work orientation using functional relationships established via human factors literature reviews

f. generating a variety of selectable output options at various levels of detail including performance effectiveness

g. simulating maintenance personnel waiting for equipment, events, and for each other

h. maintaining records of maintenance crew performance such as task duration, idle, and waiting time

Some tentatively selected values for several of the key variables are:

	Maximum Current Value
Number of Maintenance Personnel Simulated	8
Position Titles (e.g., Maintenance Mechanic, I&C Technician, Supervisor)	5
Task Duration Time (maximum in days)	2
Subtasks per Task	100
Ability/Capability Levels	7

Accordingly, the model simulates the performance of teams of up to eight persons working at any one time who may be defined into five position title categories, who perform a task which takes 48 or fewer hours and which represents any single repair/maintenance sequence performed in a nuclear power plant. Each task is represented by or segmented into no more than 100 maintenance team subtasks. The model allows a "user" seated at a computer terminal to enter parameters, to select a task for simulation, to provide other required input data (or to change existing data tables), and to initiate a computer run of the model.

5.3.3. General Model Description

The simulation process is shown generally in Fig. 5.1. The results of each run are available either at the computer's terminal display and its hard copy device and/or on the computer's high speed line printer. This makes practical a series of user-initiated runs in which the initial conditions for each subsequent run can be dependent on the results of prior runs.

Also of interest in gaining insight into the planned capabilities of the MAPPS model are features which are not currently planned for inclusion. Some of these are shown in Table

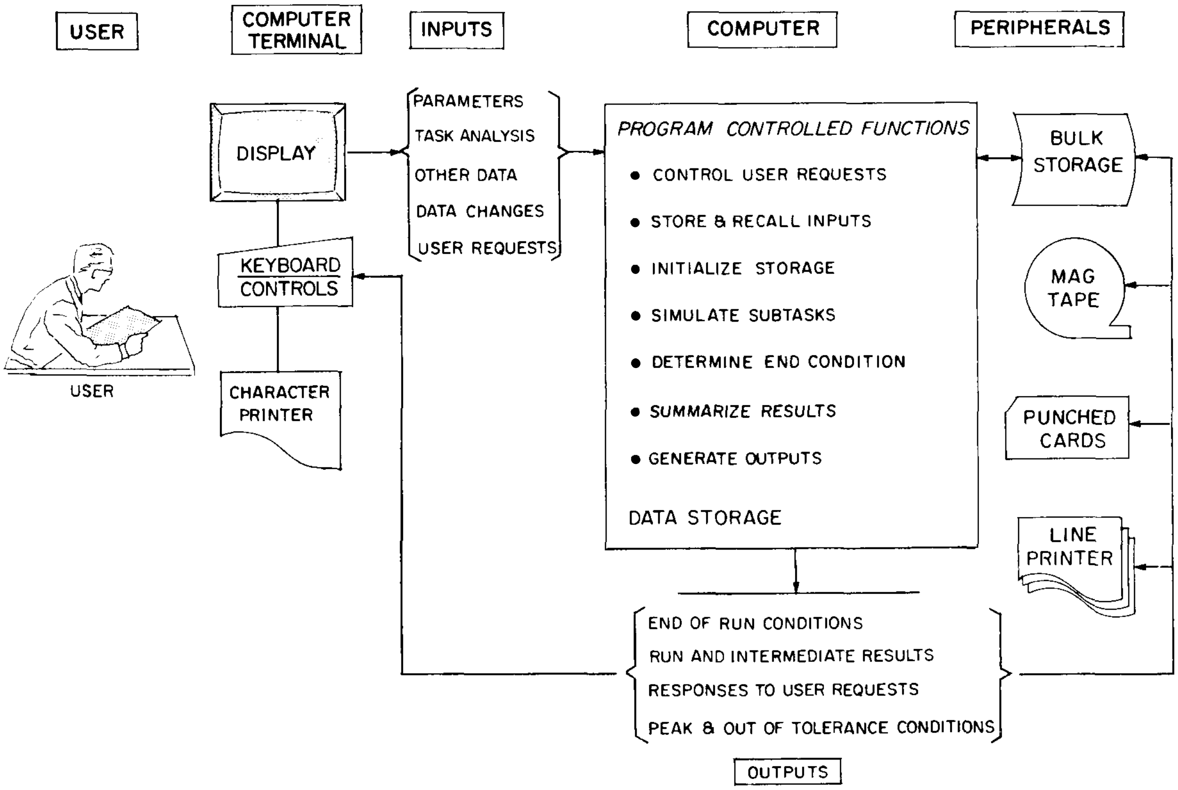

Fig. 5.1. Elements of the MAPPS Model.

5.3. Several of these features can be taken into consideration by the expedient of preparing alternate (additional) task analyses and/or team replacement data sets.

5.3.4. Global Logic Flow

Figure 5.2 shows the global logic flow of the MAPPS in a somewhat more detailed form. Each rectangular box in the figure presents a computer program module's name, number, and function. The interrelationships between modules is controlled by diamond-shaped decision boxes and circular connectives. One pass through this sequence represents completion of one run of the computer simulation model, i.e., one task simulation.

5.3.5. Model-User Interface

Consider now the user-computer interface from which a better understanding can be gained about how the model and the model's user interact. The simulation process, or interactive session, begins by the user logging-on his terminal. Following log-on, the user requests access to the MAPPS program (all user requests are made by keyboard entry) and a display similar to that shown in Fig. 5.3 appears on the terminal, with message 1 in

Table 5.3. Features Not Simulated/Calculated

- Rest or sleep periods for simulated personnel
- Malevolent behavior by maintenance personnel
- Incapacity or sickness of maintenance personnel
- Inventory levels of spare parts and expendable supplies
- Physical placement of the equipments maintained, tools, test equipment
- Failure of replacement parts
- Crosstraining

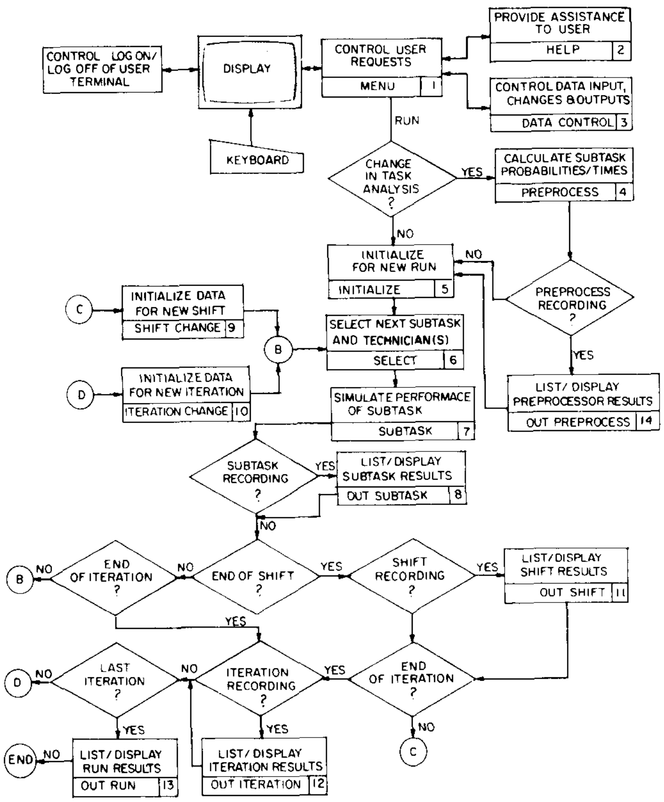

Fig. 5.2. Global Flow of the MAPPS Simulation.

the first line. This display is called the MENU. Possible messages are presented in Table 5.4 by number. The boxed-in areas in Fig. 5.3 are reserved for data entry by the user. The user enters:

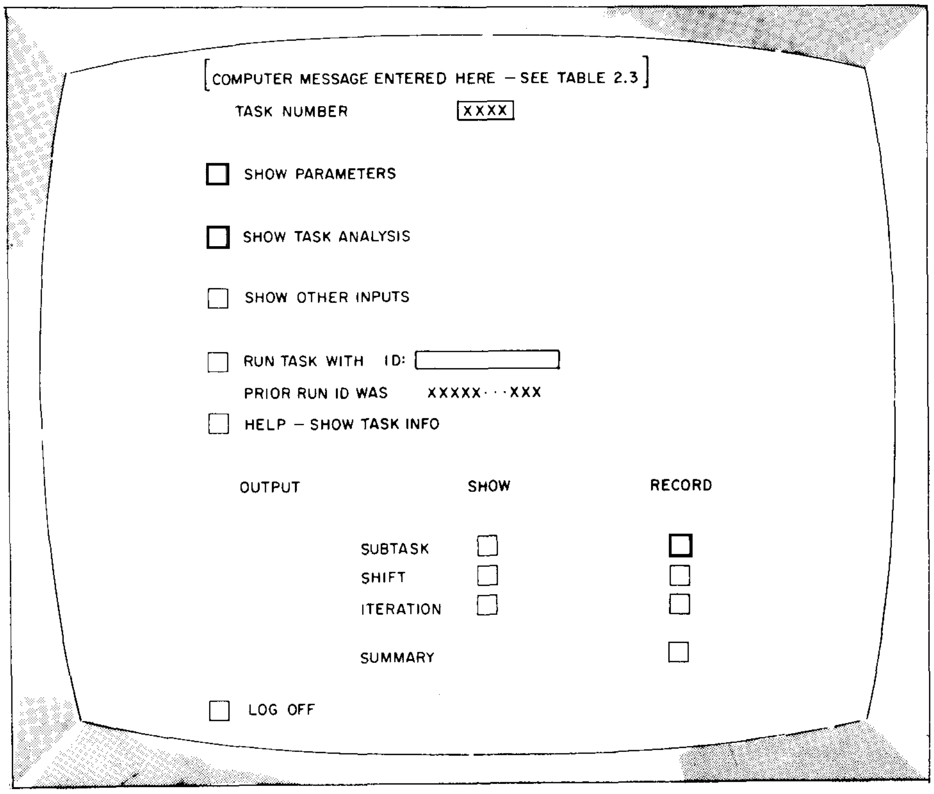

Fig. 5.3. Potential Model Menu for Terminal Display.

Table 5.4. Computer Generated Messages for MENU

1. LOG ON OK MAPPS IS READY. ENTER SELECTIONS
2. DATA ENTERED
3. DATA RECORDED
4. MENU FOLLOWS. ENTER SELECTIONS
5. YOUR ERROR IN ____________ DUE TO ____________ (a variety of error conditions will be identified)

---

1. the task number to be simulated
2. his request which can be any one of the following items:
   a. display of current parameters, task analysis, or other input values (see below)
   b. simulate the task
   c. request "help"

3. the level of detail at which output is desired — provided only if a run is requested.

Assume that the user wants to enter new input data prior to a model run. He enters any character in the "show parameters," "show task analysis," or "show other inputs" box (one at a time) and presses the TRANSMIT (or similar return) key on his keyboard. The request is processed by the MENU subroutine and the requested display is generated (see Fig. 5.4). The user then enters the data into a selected blank area of Fig. 5.4 representing individual input values and enters a character in a box at the bottom of the display to indicate his next request. His options here are either to ENTER these data (replacing any prior data item), RECORD this displayed set of data on the line printer, request SHOW of the next page of data (for task analysis only) or SHOW MENU to return the Fig. 5.3 MENU to the screen. In this way, all new input data are entered, visually verified, and/or changed. Message numbers 2 and/or 3 are displayed at the top of MENU after entry and/or recording of input data. When all data are entered to the user's satisfaction, the user can request a run by entering any character in the SHOW MENU box, entering any character in the RUN TASK box, entering task identification information (e.g., task name, date), selecting the output desired, and pressing his TRANSMIT key. For reference, the identifier used in the most recent prior run is displayed in the MENU. Output is selected by level (end of each subtask, end of shift, and end of iteration; end of run is mandatory and need not be entered). If line printer recordings of the output are desired, they may be requested at each of the four levels by entering a character in one or more of the RECORD boxes before pressing TRANSMIT. User errors are identified when an improper request is made such as data entry outside of entry areas, entry of values which are outside of "normal" ranges, requesting both SHOW PARAMETERS and RUN simultaneously.

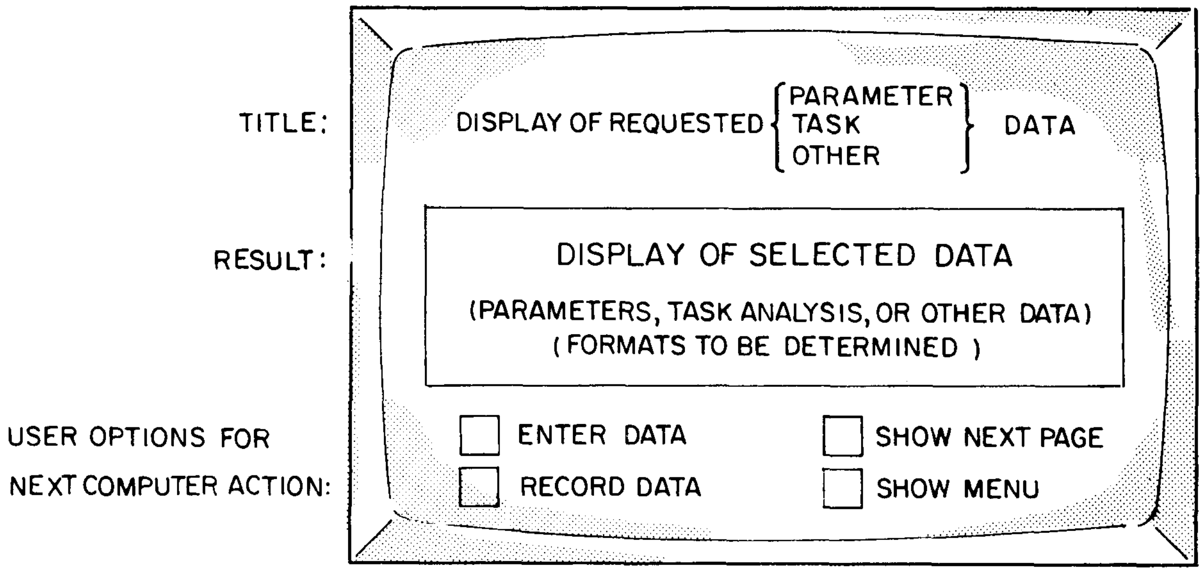

Fig. 5.4. Terminal Display When Operator Requests "SHOW."

The HELP function may be invoked by entering a character in the HELP box. A terminal display such as that shown in Fig. 5.5 is then presented giving the title, name, and data of the last MAPPS run for each task number in sequence, together with an indicator (Y=yes, N=no) as to whether or not task analysis data are stored (available) for that task.

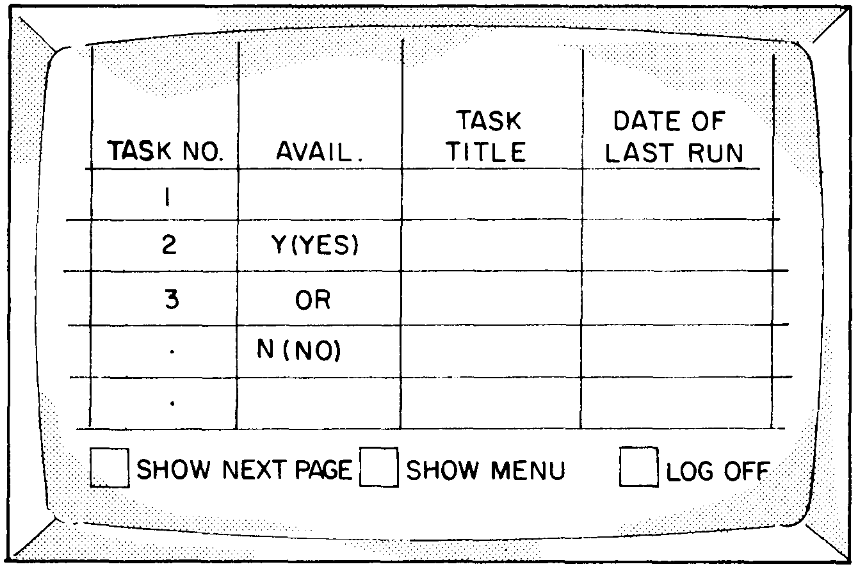

Fig. 5.5. Terminal Display When Operator Requests "HELP."

Thus the general sequence is:

1. use HELP (if needed) to select the task the user would like to simulate
2. use SHOW (if necessary) to check all input data
3. use ENTER (if necessary) to input or change any missing, incorrect, or obsolete data
4. use RUN and OUTPUT options to initiate simulations

The process repeats for as many tasks and conditions as the user desires to run, after which he logs off the terminal using procedures dependent upon his terminal type and computer facility.

5.3.6. Tentative Model Input — Parameters

A tentative list, description, and units of measure for each parameter for which input data are required prior to each run is shown in Table 5.5. The MAPPS model retains one set

Table 5.5. Tentative Model Parameters

- Temperature at place of maintenance (°F) (Default = 70°)
- Average ability levels, A, for each simulated maintenance person. Seven (0-7) values for each man, M, or type of man, P. (Default = 3 for all P values of A)*
- Stress threshold (Default = 2.5)*
- Accoustic noise level at place of maintenance (db) (Default = 0)
- Radiation dose absorbed by average team member prior to start of task (RADS) (Default = 0)
- Radiation level during task at place of maintenance (RADS/HR) (Default = 0)
- Time worked prior to start of task (average operator) (Hrs) (Default = 0)*
- Number of iterations, N, of task per simulation run (Default = 10)
- Leader's aspiration or expectation [0 (low) to 1 (high)]
- Aspiration of each team member [0 (low) to 1 (high)] (Default = same as leaders)*
- Human factors maintenance situation [0 (poor) to 1 (excellent)]; includes lighting, crowding, accessibility (Default = 0.5)*
- Time limit of maintenance task performance (Hours)
- Importance of achieving the time limit.

  [0 (unimportant) to 1 (essential)] (Default = 0.5)
- Limit value for "essentiality" over which a subtask may never be ignored (Default = 0)

*Average applicable to all maintenance personnel for the run.

(the current run set) of parameter values. Accordingly, only data representing changes from the most recent parameter set need be entered. Those items, identified by an asterisk in Table 5.5, may be entered once (e.g., an average applicable to all maintenance personnel for the run). Alternately, at the users option, individual values may be entered for each maintenance Man, M, to be simulated.

Default values are given for some parameters in Table 5.5. If the user does not provide a value for a model parameter, the MAPPS model will substitute the default value.

The time limit for the maintenance task was included in the parameter list (Table 5.5) because in the maintenance of some systems (e.g., safety systems) down time minimization is quite essential. In order to have the potential to place little or no emphasis on this time limit for some maintenance tasks, the "Importance of Time Limit" parameter is allowed to vary.

The parameter for the Number of Task Iterations per simulation is included since several of the model's variables, as well as other elements in the model, are implemented by the use of Monte Carlo techniques. These techniques utilize random numbers to select from probability distributions in order to determine step-by-step function values to be used at a particular place in the calculation. They are also used to select alternative courses of action with predetermined probabilities and to select individual initial parameter values for each simulated maintenance technician, based on average input parameters provided by the user. This nonanalytical approach results in distributions of results whose average error is dependent on the number of samples (run iterations) taken. Accordingly, it is assumed that a given result will have some inherent sampling inaccuracy (standard error) but that the error is reduced by larger numbers of simulation iterations.

Technician ability on seven ability factors are indicated by the user to the computer program by the identifier A. Table 5.6 shows the code assignments for A as well as the codes for kinds of subtasks (K), job titles of personnel simulated (P), types of protective clothing (C), and other key model indices.

5.3.7. Tentative Model Input — Task Analysis Data

Each simulated task is analyzed and subdivided into a number of subtasks (I). Each subtask of every task is quantified by providing the data elements specified in Table 5.7. Here, again, the table provides units of measure, input value ranges, and default values as appropriate. A single list is prepared; not a separate list for each man.

Task analysis data are entered by the user at his terminal (general form of Fig. 5.4) after making the SHOW TASK ANALYSIS request (Fig. 5.3).

Each task analysis is prepared assuming a fixed number (less than or equal to eight) of maintenance personnel. At times of specified personnel replacements (shift changes), other personnel are assumed available (with the limit that no more than eight are permitted at one time) and that at all times the number of men working will be no greater than the number assigned at the start of the task. Initial conditions at shift change (e.g., aspiration, prior radiation exposure level) will be calculated using the same input parameters that are used for each corresponding man at the start of the task (first shift).

The essentiality value is a measure of the criticality of the subtask for completion of the task. A subtask having a low essentiality value may be skipped during simulations in which there is high time stress.

Protective clothing change data are required only for subtasks just prior to which one or more of the assigned men has a change in clothing assignment. The definition of protective clothing assignments at the start of the task is given later under "other data."

5.3.8. Tentative Model Input — Other Data

The remainder of the required input data are grouped into the category of "other data" and are itemized in Table 5.8.

Table 5.6. **Tentative Model Indices and Nomenclature**

Index	Range	
M	1-8	Identifier for maintenance personnel (Man No.)
S	1-24	Identifier for shifts
I	1-100	Identifier for subtasks
C	1-6	Code for type of protective clothing 0-None 1-Minimum 2-Full
P	1-6	Code for personnel category 1-Mechanic; 4-Electrical 2-Instrument and Control; 5-Quality Control 3-Supervisor
K	1-8	Code for kind of subtask 1-Motor; 5-Clothing change 2-Decision; 6-Communications 3-Perceptual; 7- } to be determined 4-Mental; 8- } to be determined
A	1-7	Identifier for ability taxonomy 1-Visual speed, accuracy, and recognition 2-Gross motor coordination 3-Fine manual dexterity 4-Strength and stamina 5-Cognition 6-Memory 7-Problem solving
T	1-3	Code for type of communication subtask 1-Gives orders 2-Makes suggestions 3-Requests information

5.3.9. Functional Relationships and Variables

Functional relationships for the selected model variables will be defined in specific mathematical and/or logical form during the development phase of the program, as described later in this report. The present definition is in the form of principal functional dependencies to be utilized (Table 5.9) and those involved in the end of iteration calculations (Table 5.10).

Consider now the group concept. For each subtask, one or more men, M, are identified as required for subtask completion. The collection of such men is called a team.

Table 5.7. Tentative List of Task Analysis Data

The following data are provided sequentially for each subtask, I, i.e., one sequential subtask list:

- subtask No., I, an integer from 1 to 100
- essentiality from 0 (low) to 1 (high) (Default = 1)
- next subtask number, integer from 1 to 100
  - a. if successful, default = I + 1
  - b. if failed, default = I (repeat)
- probability of supervisor detection of error (0 to 1)
- kind of subtask (K); see Table 5.6
- protective clothing change data, an optional input for up to 8 men:
  - a. man no., M, changing; an integer from 1 to 8
  - b. type of clothing changed into, C, see Table 5.6 (Default = 0, 0) change of clothing data are effective at start of subtask
- time until which maintenance man must wait before beginning this subtask (hours after start of task). (Default = 0)
- personnel assigned to this subtask; up to 8 values of M
- effect of failure of this subtask on:
  - a. safety level from 0 (low) to 1 (high). (Default = 0)
  - b. radiation level (RADS/hr. change). (Default = 0)
- subtask precedence indicator specifying whether this subtask:
  - a. can begin at any time, i.e., each man can start at a different time
  - b. must begin only after all required men are available at the same time. (Default = b)
- indicator of whether or not a quality control check must be made after the subtask

  (Y = Yes, N = No) (Default = N)
- indicator of any subtask which must be completed before this subtask can be started (subtask precedence number)
- shift change data table:

Shift No.	Applicable to Personnel No. Enter 1 = Yes or 0 = No	Shift Change Conditions Subtask No.	Time (Hrs.)
	1 2 8		
1			
2			
3			
.			
.			
.			
S			

For each shift, S, identify the man (or men) who are to be replaced and the conditions of shift change. Condition for change is either subtask no., time since task start, or both (whichever occurs first).

Table 5.8. Other Tentative Input Data

- Task Number — indicating which maintenance task is to be simulated, 0 to 9999 (see Figure 5.3)
- Run ID designating a specific task run. (See Figure 5.3); could include, to user option, one or more of the following:

  a. task
  b. run number
  c. date
  d. user name

- User request for service as shown in Menu, Figures 5.3 and 5.4:

a. show parameters	g. enter changes
b. show task analysis	h. show next page
c. show other inputs	i. show menu (default)
d. help	j. record data
e. output requests (7)	k. log off
f. run task	

- Personnel category, P(M) and initial clothing assignment C(M) for each M

Man M	Code for Category of Personnel P(M)	Initial Clothing Code C(M)
1		
2	category, no. of personnel assigned see Table 5.6	type of clothing worn at start of task
up to 8		

- Limit in RADS over which output is flagged.
- Limits in { Degrees F / Time in protective clothing } after/over which outputs are flagged.

Figure 5.6 shows a time-line sequence to illustrate the team member and other time dependent concepts involved in subtask simulation. In this illustrative eight-man task, subtasks 1 and 4 are performed by men 1 and 2; subtask 2 is performed by men 3, 4, and 5; subtask 5 is performed by a three man group composed of men 2, 3, and 4, etc. Figure 5.6 illustrates several model situations:

1. Subtask precedence indicators (see Table 5.7) can be of type A (subtask begins anytime) e.g., subtasks 1, 8, 9 or of type B (subtask begins only after all men are available) e.g., subtasks 5, 6. For type A, not all men need start the subtask at the same time.

Table 5.9 Tentative Model Variables Calculated at the Subtask Level

- Radiation absorbed dose (M) = f (prior dose, subtask duration, radiation level parameter, fail/succeed result, effect of failure on radiation level)
- Performance (M) = f (no. of subtask failures, successes, undetected errors)
- Skip subtask = f (essentiality, time stress, time limit, importance parameter)
- Work orientation (M) = f (aspiration, time spent in failure/repetition/recovery, leader's aspiration)
- Time stress = f (time limit, subtask duration, work remaining, importance of time limit parameter)
- Total stress = f (time stress, radiation exposure level, temperature, human factors maintenance situation, stress threshold)
- Subtask success = f (probability of subtask success, total stress, work orientation (M), clothing indicators)
- Probability of subtask success = f (ability levels of personnel, ability levels required)
- Subtask duration = f (kind of task, human factors maintenance situation, average subtask duration, stress, stress thresholds, fatigue, noise level, temperature)
- Average subtask duration = log normal transform of time
- Fatigue (M) = f (hours since shift start, clothing code)
- Aspiration change = f (stress, aspiration difference (worker/supervisor))

Table 5.10. Tentative Model Variables Calculated at the End of Each Iteration

Risk weight	=	f (undetected errors, team safety level, risk rate parameter)
Team safety level	=	f (crew absorbed dose team, undetected errors, human factors situation, team fatigue, temperature)
Team absorbed dose	=	$\sum_{M}$ Radiation absorbed dose (M)/No. of men
Team fatigue	=	$\sum_{M}$ Fatigue (M)/No. of men
Team undetected errors	=	$\sum_{M}$ Final undetected errors (M)
Effectiveness	=	f (risk weight, task time overrun, crew performance)
Team performance	=	$\sum_{M}$ Performance (M)/No. of men
Final team stress	=	$\sum_{M}$ Final team stress (M)/No. of men
Task time overrun	=	f (task duration, time limit)

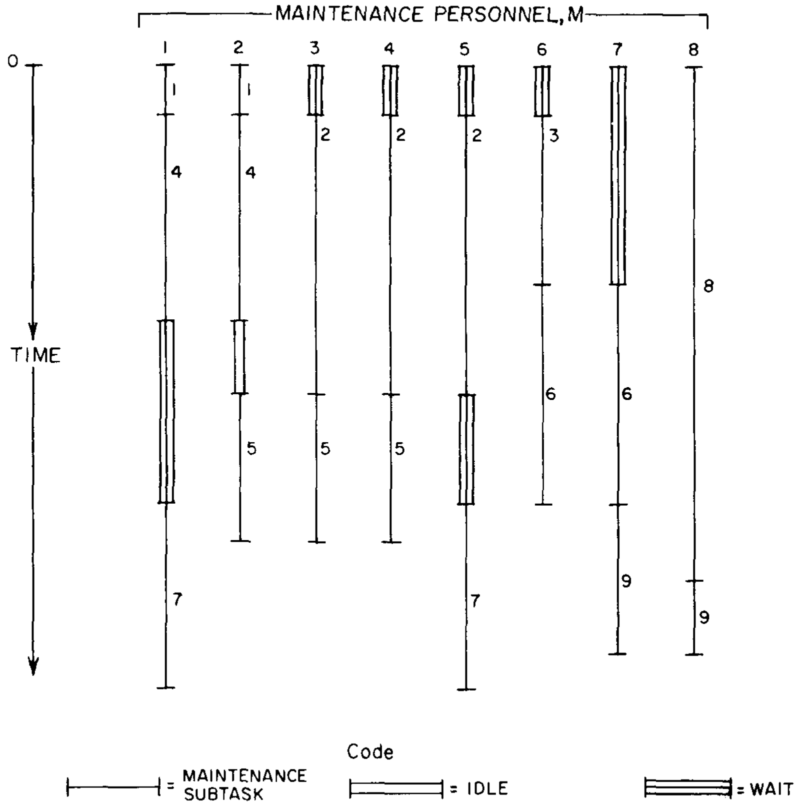

Fig. 5.6. Example of Subtask Flow Interdependencies.

2. An "idle" is required as a subtask precedence indicator (type B) when all men are not available at the same time (e.g., subtask 5). These are indicated by double vertical lines.

3. The subtask precedence number (Table 5.7) indicates, for example, that subtask 3 cannot begin until subtask 1 has been completed; also, subtask 7 waits for number 6. In these cases a WAIT is induced as indicated by a triple vertical line.

5.3.10. Tentatively Planned Model Output

The numerical results generated as a result of running the MAPPS model include the four categories shown in Table 5.11. Output of these results plus recording of inputs is available as shown below:

Table 5.11. Potential Model Outputs

Title for all Outputs: Task No., Run ID, Date, Type of Report

Output	Output Category			
	Subtask	Shift	Iteration	Summary
Subtask no.	X			
Start time, end time	X			
duration	of subtask	of shift	of iteration	*
shift no.		X		
iteration no.			X	
result (success/failure)	X			
For each man:				
current radiation absorbed dose	X	X	total	*
performance		X	X	*
fatigue	X			
time worked, idle, wait	X		total	
work orientation	X		X	
Number of men	in team		in team	in team
For Entire Maintenance Team:				
no. of subtasks completed, failed			X	*
stress			X	*
team safety level			X	*
radiation absorbed dose			average & max	*(1)
subtasks ignored			X	*
fatigue			X	*
risk weight			X	*
effectiveness			X	(2)
time spent in:				
maintenance			X	*
repeats			X	*
idle			X	*
wait			X	*
task duration (MTTR)			X	*
no. of iterations				X
undetected errors			X	*
% iterations under time limit			X	*
tasks failed			(2)	(2)

*Average value calculated over all iterations in the run (1) number of men over limit (2) frequency distribution.

Outputs	Output Availability At User Terminal	At Line Printer
Subtask		X
Shift		X
Iteration	X	X
Summary	X	X
Parameters	X	X
Task Analysis	X	X
Other Inputs	X	X

The detailed output generated at the completion of each subtask, and shift is expected to be useful primarily for model testing and user training. Examination of end-of-iteration results allows the user to determine the extent of fluctuations and variations in model variables and outputs over the several task iterations. Of greatest utility are the summary results generated at the end of simulation of each N iteration run. As indicated in Table 5.11, summary output includes averages over the N iterations for all key results. It is also planned to provide key summary data on the computer terminal in such a way that the hardcopy recording of a single page will capture the title information, summary task results, and the key parameter values. In this way, the essence of each run maybe recorded on a single page for archival purposes.

In general, the selection of likely outputs, shown in Table 5.11, is presented as a maximum list. The greater the extent of output variety and richness, the more extensive are the required calculations to generate them. Since it is desired that MAPPS be useful as a terminal-driven program, it is also recognized that response times to user requests must be reasonable. Accordingly, some reductions are planned in the output list shown with the objective of running tasks of average lengths — say 50 to 60 subtasks — over 100 iterations in an elapsed time period of 2 to 5 minutes.

5.4. Program Plan

Section 5.3 of this report defined some of the tentative global features of the MAPPS model. Section 5.4 presents a plan to develop, validate, and disseminate this model. Appendix E of this report provides additional details concerning this plan.

5.4.1. Program Plan Overview

In order to place the total plan in perspective, consider Fig. 5.7 which ties together the various phases in the life of any computer model. Figure 5.7, which is read from bottom to top, shows the major steps (large rectangles) representing efforts from conceptualization and model requirements derivation through the situation in which the model can be considered adequate for user decision-aiding and eventually for decision-making. The figure also suggests the various "states" of a model (in shaded ovals) over time. For example, a model which is programmed and debugged in accordance with the stated model specifications (requirements) is said to enter a state called "testable;" after sensitivity testing, and implementation of corrections, the model is said to be "reasonable." The various types of

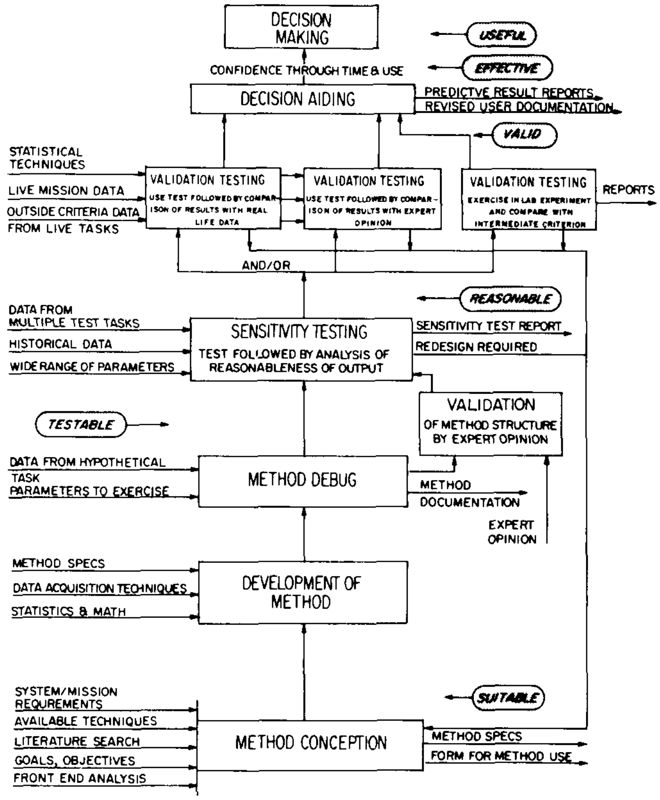

Fig. 5.7. Sequence of Model Development.

data and information required as inputs to each phase are shown entering from the left with the resulting end products exiting to the right.

Considering the total effort involved in the life of a computer model, as displayed generally in Fig. 5.7, this report is equivalent to the method specifications, and the work of method conception is now completed.

5.4.2. General Program Plan

The program plan for the MAPPS simulation model spans a 38-month period of time and consists of three distinct phases. The three phases are: the development phase, the validation phase, and the dissemination phase. Each of these phases will be discussed in subsequent subsections.

Two model releases are planned within this program. The first will be a debugged, sensitivity-treated version of the model which will be released to the NRC (only) following the completion of the development phase. The second release will be subsequent to the completion of the validation phase and will be a validated version of the model. The release will be to the NRC and the public.

Development, validation, and dissemination of the MAPPS model will also be subject to peer review. Several peer review meetings have been planned throughout the program. Their purpose is to provide to the project team an independent evaluation of the effort within the various phases of the program, to make suggestions for improvement, to identify relevant aspects of the maintenance context not addressed (if any), and to provide general guidance for development, testing, validation, and dissemination of MAPPS.

5.4.3. Model Development Phase

The model development phase of the program will be concerned primarily with the formulation and development of a computer simulation model to predict the reliability of nuclear power plant maintenance personnel in the performance of maintenance tasks.

The general logic of the model will be developed from appropriate psychosocial, maintenance, reliability, and probabilistic theories. The general model logic will identify interdependencies between modules that address specific parameters within the model such as stress, fatigue, the effects of temperature, etc. Explicit functional relationships between model variables will be developed and set into the simulation framework of the model. The model will be programmed, debugged, and sensitivity tested across the ranges of all input parameters. In addition, each of the modules will be tested individually for reasonability of output.

Task analyses for a set of sample tasks will be accomplished to provide data for the calibration of the model. Calibration of the model will be complete when results generated by the model seem "reasonable." Once the model has been calibrated, the state-of-the-model will be kept intact for validation purposes. The debugged/sensitivity-tested/calibrated version of the model (Version 1) will be released to NRC upon completion of model calibration. Further details concerning the development phase may be found in Appendix E.

5.4.4. Model Validation Phase

Model validation will be accomplished according to a validation plan that will be developed early in this phase, and will be addressing Version 1 of the model. The plan will consist of a validation strategy which will be the end result of the consideration of a number of validation approaches. In addition, the plan will include a specification of the tasks to be addressed during this phase of the program. Selection of the validation approach and the validation tasks will be made by considering various types of criteria including practicality acceptance and usefulness.

Implementation of the approach will yield two primary sets of information. The first will be a comprehensive profile of maintenance performance with respect to each of the selected tasks. The second will be a complete set of team variables for input into the MAPPS model.

The validation phase of the program will include the performance of task analyses for the selected validation tasks. The task analysis data, along with the set of team variables generated from implementation of the selected approach will be used as input data to

MAPPS. The model will then be run for each of the selected tasks to produce a simulated profile of maintenance performance. Statistical comparisons between the two profiles will be performed to determine if statistical differences exist. Favorable results obtained from comparing the maintenance profiles will demonstrate the model's external validity. The model's internal validity will also be addressed by examining the significance of the module to model and the module to module correlations.

Upon completion of this phase of the program, a validated version of the model will be released to NRC and the public along with recommendations for possible model improvements. Further details concerning the validation Phase may be found in Appendix E.

5.4.5. Model Dissemination Phase

The objective of the model dissemination phase is to effectively transfer to the NRC and to other potential users, the methodology developed within this program.

Transfer of the methodology will be done according to a dissemination plan developed early during this phase. The plan will describe a comprehensive workshop that will have as its goal the effective transfer and efficient implementation of the model by potential users.

Potential users from NRC and other organizations will be contacted and provided with preliminary workshop materials. The workshop will provide the opportunity for each participant to acquire a detailed working knowledge of the model, requirements for running the model, and an ability to interpret the metrics produced by the model. Hands-on experience will be emphasized to encourage familiarity. A portion of the workshop will be aimed toward conducting successful task analyses for the generation of input data for the model. Further details of the dissemination phase may be found in Appendix E.

5.5. Program Plan Summary

Chapter 5 of this report has presented a general description of the MAPPS model and a detailed plan for its development, validation, and dissemination. The total time span for this program is 38 months and will result in a fully validated simulation model capable of providing measures of expected performance of nuclear power plant maintenance personnel performing maintenance tasks. At the culmination of this program the model will have been transferred to the NRC and other potential users, and a workshop for effective utilization of the model will have been provided. Appendix E of this report provides a detailed description of the program plan and estimated schedule for its implementation.

6. SUMMARY OF THE REPORT

The front-end analysis for the nuclear power plant maintenance personnel reliability model consisted of three primary tasks. These were: a front-end user survey, a literature review of human behavioral methodologies, and the development of a program plan for the proposed model.

The front-end user survey was conducted in order to determine the need for such a model and at the same time determine the general content and focus of the model. Results of the survey indicated that a computerized methodology that could produce estimates of various measures of performance would be very useful in a number of contexts. The survey also identified the types of information that would be most useful and thus aided in determining the general nature of the model.

The literature review of human behavioral methodologies was conducted in order to gain an insight into the state-of-the-art of existing predictive methodologies. The primary objective of this portion of the front-end analysis was to evaluate existing methodologies in order to select the type of methodology to be developed within this program. A general evaluation of the methodologies was carried out with respect to certain desirable traits and model characteristics, and simulation modeling was chosen as the type of methodology to be developed. Results from the front-end survey, indicating the desired informational outputs of a computerized methodology seemed to support the choice of simulation modeling.

The program plan for the MAPPS model was developed to provide details concerning the development, validation, and dissemination of the simulation methodology. The 38-month program includes two releases of the model. Version 1, which is the debugged/sensitivity tested/calibrated version, will be released to the NRC following the development phase of the program. Version 2, which is the validated version, will be released to NRC and any other organization designated to receive the model, following the validation phase. The dissemination phase of the program will include the conduction of a workshop that will emphasize effective utilization of the model and proper interpretation of the model's output.

The program as described in Chapter 5 and Appendix E will be subject to peer review. Input to the program from subject matter experts participating in the review will provide an independent evaluation of program efforts, suggest improvements, identify relevant aspects of the maintenance context that may not have been addressed (if any), and in general provide expert guidance for the development, validation, and dissemination phases.

This report has also presented details concerning the functional requirements of a maintenance model (Chapter 4) as well as a general list of potential uses and users of the model. The development of MAPPS is being done primarily to support the human reliability analyses portion of NRC's PRA programs. The model will be a valuable source of data and information for NRC concerning the maintenance area and will significantly contribute a much needed data source for a wide variety of other users requiring human reliability data for nuclear power plant maintenance.

7. REFERENCES

1. J. P. Finnegan, T. W. Rettig, and C. A. Rau, Jr., "The Role of Personnel Errors in Power Plant Equipment Reliability," Failure Analysis Associates, Palo Alto, CA, 1979.
2. R. E. Hall, P. K. Samanta, and A. L. Swoboda, "Sensitivity of Risk Parameters to Human Errors in Reactor Safety Study for a PWR," NUREG/CR-1879, Brookhaven National Laboratory, Upton, NY, 1981.
3. IEEE Task Group, "Guide for Human Factors Engineering Requirements for Systems, Equipment, and Facilities of Nuclear Power Generating Stations," Draft Report, 1981.
4. W. J. Lukas, "Initial Quantification of Human Errors Associated With Reactor Safety System Components in Licensed Nuclear Power Plants," R. E. Hall (Ed.), Brookhaven National Laboratory, Upton, NY, 1981.
5. A. D. Swain, "Handbook of Human Reliability Analysis With Emphasis on Nuclear Power Plant Applications," NUREG/CR-1278, H. E. Guttman (Ed.), Sandia National Laboratories, Albuquerque, NM, 1980.
6. H. E. Price, R. E. Maisano, and H. P. Van Cott, "The Allocation of Functions in Man-Machine Systems: A Perspective and Literature Review," An Interim Report prepared by BioTechnology, Inc. for the Oak Ridge National Laboratory under Subcontract No. 9027, September 1981.
7. D. Meister, *Human Factors: Theory and Practice*, pp. 21-56, John Wiley & Sons, Inc., New York, 1971.
8. L. V. Rigby, "The Sandia Human Error Rate Bank (SHERB)," SC-R-67-1150, Sandia National Laboratories, Albuquerque, NM, 1967.
9. D. Meister, "Analysis of Human Initiated Equipment Failures During Category 1 Testing, OSTF-1," REL R-054, General Dynamics/Astronautics, San Diego, CA, 1961.
10. Electric Power Research Institute, "Human Factors Review of Power Plant Maintainability," EPRI NP-1567, February 1981.
11. J. D. Davis and M. W. Levine, "A Model for the Control of Ingestion," *Psychological Review* **84**, 379-412 (1977).
12. R. R. Kramer, "A Simulation of the Human Shoulder," U.S. Army Human Engineering Laboratory, Aberdeen Proving Grounds, MD, 1975.
13. S. Penrod and R. Hastic, "Models of Jury Decision Making: A Critical Review," *Psychological Bulletin* **86**, 462-492 (1979).
14. D. E. Embrey, Personal Communication, 1981.
15. A. D. Swain, Personal Communication, 1982.
16. D. Meister, Personal Communication, 1982.
17. D. Meister, "Comparative Analysis of Human Reliability Models," Contract N00024-71-C-1257, Bunker-Ramo Corp., Westlake Village, CA, 1971.

18. D. Meister, "A Critical Review of Human Performance Reliability Predictive," *IEEE Transactions on Reliability* **R-22(3)**, 116-123 (1973).

19. D. E. Embrey, "Human Reliability in Complex Systems: An Overview," NCSR R-10, National Centre of Systems Reliability, July 1976.

20. H. A. Simon and A. Newell, "Human Problem Solving: The State of the Theory in 1970," *American Psychologist* **26**, 145-159 (1971).

21. A. Newell and H. A. Simon, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, NJ, 1972.

22. A. I. Siegel, J. J. Wolf, and A. R. Williams, "Computer Simulation of Human Performance in Electronic Imagery Systems," Applied Psychological Services, Inc., Wayne, PA, 1978.

23. S. J. Munger, R. W. Smith, and D. Payne, "An Index of Electronic Equipment Operability: Data Store," American Institute for Research, Pittsburg, PA, 1962.

24. D. Payne and J. W. Altman, "An Index of Electronic Equipment Operability," AIR-C-43-1162, American Institute for Research, 1962.

25. A. D. Swain, "A Method for Performing a Human Factors Reliability Analysis," Monograph SCB-685, Sandia National Laboratories, Albuquerque, NM, August 1963.

26. A. D. Swain, "Field Calibrated Simulation," *Proceedings of the Symposium on Human Performance Quantification in Systems Effectiveness*, Naval Material Command and the National Academy of Engineering, Washington, D.C., January 1967.

27. A. D. Swain, "Overview and Status of Human Factors Reliability Analysis," *Proceedings of the 8th Annual Reliability and Maintainability Conference*, American Institute of Aeronautics and Astronautics, Denver, Colorado, July 1969.

28. A. D. Swain, "Development of a Human Error Rate Data Bank," *Proceedings of U.S. Navy Human Reliability Workshop, July 22-23, 1970*, J. P. Jenkins (Ed.), Naval Ship Systems Command, Office of Naval Research, and Naval Air Development Center, Washington, D.C., February 1971.

29. A. D. Swain, "Human Factors Associated With Prescribed Action Links," SAND 74-0051, Sandia National Laboratories, Albuquerque, NM, July 1974.

30. U.S. Nuclear Regulatory Commission, "Reactor Safety Study in U.S. Commercial Nuclear Power Plants (WASH-1400)," NUREG-75/014, Main Report, Appendix III — Failure Data, Washington, D.C., October 1975.

31. A. D. Swain, "Human Reliability Assessment in Nuclear Reactor Plants," Monograph SCR-69-1236, Sandia National Laboratories, Albuquerque, NM, April 1969.

32. R. E. Blanchard, M. B. Mitchell, and R. L. Smith, *Likelihood of Accomplishment Scale for a Selection of Man-Machine Activities*, Dunlap and Associates, Inc., Santa Monica, CA, 1966.

33. R. L. Smith, R. A. Westland, and R. E. Blanchard, "Technique for Establishing Personnel Performance Standards (TEPPS), Technical Manual," Report PTB-70-5, Vols. 1 and 2, Personnel Research Division, Bureau of Naval Personnel, Washington, D.C., December 1969.

34. R. L. Smith, R. A. Westland, and R. E. Blanchard, "Technique for Establishing Personnel Performance Standards (TEPPS), Results of Navy User Test," Report PTB-70-5, Vols. 3, Personnel Research Division, Bureau of Naval Personnel, Washington, D.C., December 1969.

35. A. B. Pontecorvo, "A Method of Predicting Human Reliability," *Annals of Reliability and Maintenance* **4**, 337-342, 4th Annual Reliability and Maintainability Conference (1965).

36. A. I. Siegel and J. J. Wolf, "Digital Behavioral Simulation — State-of-the-Art and Implications," Report prepared for the U.S. Army Research Institute for the Behavioral and Social Sciences, Applied Psychological Services, Inc., Alexandria, VA, March 26, 1981.

37. A. I. Siegel and P. J. Federman, "Prediction of Human Reliability: Part 1: Development and Test of a Human Reliability Predictive Technique for Application in Electronic Maintainability Prediction," Applied Psychological Services, Inc., Wayne, PA, 1971.

38. A. I. Siegel and J. J. Wolf, *Man-Machine Simulation Models*, John Wiley & Sons, Inc., New York, 1969.

39. A. I. Siegel, J. J. Wolf, and A. R. Williams, "A Model for Predicting Integrated Man-Machine System Reliability: II. Model Validation," Applied Psychological Services, Inc., Wayne, PA, 1976.

40. A. I. Siegel, J. J. Wolf, and M. R. Lautman, "A Model for Predicting Integrated Man-Machine Systems Reliability," Applied Psychological Services, Inc., Wayne, PA, 1974.

41. A. I. Siegel, J. J. Wolf, and J. Cosentino, "Digital Simulation of the Performance of Intermediate Size Crews: Application and Validation of a Model for Crew Simulation," Applied Psychological Services, Inc., Wayne, PA, 1971.

42. NAVSEA 1977, "Human Reliability Prediction System Users Manual," Department of the Navy, Sea Systems Command, Washington, D.C., December 1977.

43. A. I. Siegel, J. J. Wolf, J. D. Barick, and W. Michle, "Digital Simulation of Submarine Crew Performance: 1. Logic of a Psychosocial 'Model' for Digitally Simulating Crew Performance," Applied Psychological Services, Inc., Wayne, PA, 1964.

44. D. J. Seifert and G. P. Chubb, "SAINT: A Combined Simulation Language for Modeling Large Complex Systems," AMRL-TR-78-48, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, October 1978.

45. D. B. Wortman, S. D. Duket, D. J. Seifert, R. L. Hann, and G. P. Chubb, "Simulation Using SAINT: A User-Oriented Instruction Manual," AMRL-TR-77-61, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, July 1978.

46. D. B. Wortman, S. D. Duket, D. J. Seifert, R. L. Hann, and G. P. Chubb, "The SAINT Users' Manual," AMRL-TR-77-62, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, June 1978.

47. S. D. Duket, D. B. Wortman, D. J. Seifert, R. L. Hann, and G. P. Chubb, "Documentation for the SAINT Simulation Program," AMRL-TR-77-63, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, June 1978.

48. S. D. Duket, D. B. Wortman, D. J. Seifert, R. L. Hann, and G. P. Chubb, "Analyzing SAINT Output Using SPSS," AMRL-TR-77-64, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, July 1978.

49. A. I. Siegel and J. J. Wolf, *Man-Machine Simulation Models: Performance and Psychological Interaction*, John Wiley & Sons, Inc., New York, 1967.

50. A. A. B. Pritsker, *The GASP-IV Simulation Language*, John Wiley & Sons, Inc., New York, 1974.

51. A. I. Siegel, W. D. Bartter, and F. F. Kopstein, "Job Analysis of Maintenance Mechanic Position for the Nuclear Power Plant Maintenance Personnel Reliability Model," NUREG/CR-2670, ORNL/TM-8301 (June 1982).

52. W. D. Bartter, A. I. Siegel, and P. J. Federman, "Job Analysis of the Maintenance Supervisor and Instrument and Control Supervisor Position for the Nuclear Power Plant Maintenance Personnel Reliability Model," NUREG/CR-2668, ORNL/TM-8299 (November 1982).

APPENDIX A

INITIAL INTERVIEW CONTENT

Items for Interview

Maintenance Technician

1. What is the maintenance personnel structure?
2. What are job titles at this place?
3. Are job titles standardized? By what other job titles are the same jobs known?
4. Are job descriptions available?
5. Is there a job list available and a volume of procedures?
6. Are any maintenance texts available?
7. Maintenance activities--who decides what is done, when, and for how long?
8. What are the types of functions performed?
9. What is typical training of technicians? OJT?
10. Describe tasks in a typical day from start of work until going home? Elaborate on task list?
11. How is communication maintained during maintenance?
12. Is there direct supervision? Work in teams? As individuals?
13. Are time limits placed on individual activities?
14. Sources of error in maintenance activities?
15. Is there a list of the records/logs you complete?
16. Elaborate the task list with as many activities as you can?
17. How is work segmented, i.e., are there specialists for different parts of plant? Describe.
18. Design of equipment for maintenance--critical incidents?

Instrument and Control Technician

Same questions as for maintenance technician.

Foreman

1. In the case of contracted maintenance, what are the in-house tasks?
2. How is contractor maintenance decided on? (Rationales for delays.)
3. Interactions among various maintenance activities vis-a-vis scheduled maintenance? Preventive maintenance?
4. Task list availability?
5. Procedures availability?
6. Supervisory problems?
7. Organizational chart of sections?
8. His job title? Job title of others like him?
9. Training for a foreman? Training in supervisory methods?
10. List tasks (activities) performed in a typical day?
11. Expand list?
12. What logs/records/status boards are maintained?
13. What reports are reviewed?
14. Who does scheduling? How? Time limits? Cost analysis? Is activity time paced? Event paced?
15. Have any LERs been submitted?
16. Is there a procedure available for every maintenance activity?
17. Who evaluates adequacy of maintenance? Need for maintenance?
18. Environmental conditions during maintenance; design of equipment for maintenance?
19. Degree of technician initiative involved?
20. How many technicians perform each maintenance activity?

Maintenance Manager

1. How important is predictive information about:

    a. time by proficiency of operator
    b. probability of success
    c. optimum maintenance team composition
    d. effect of team morale on time, output
    e. which subtasks are most likely to be failed
    f. where intellective load is highest
    g. where psychomotor load is highest
    h. effects of training on proficiency
    i. overall safety index
    j. ability to do task under stress
    k. function allocation trade-offs
    l. training requirements

2. What factors influence maintenance effectiveness, e.g., training, stress, environment?
3. What type of human performance reliability information is most needed? Why?
4. Who would use such information?
5. How would such information be helpful?
6. What type of maintenance information is most needed from a predictive method?

APPENDIX B

STRUCTURED INTERVIEW CONTENT

Applied Psychological Services, Inc.
Science Center
Wayne, PA 19087

CONDITIONS SURVEY

The Applied Psychological Services under contract with the Nuclear Regulatory Commission and working through the Oak Ridge National Laboratory is producing a computer oriented technique for analyzing the performance of nuclear power plant maintenance personnel. This includes mechanics, I and C, and electrical personnel. The technique will describe how various personnel, environmental, and equipment design factors contribute to such considerations as cost/effectiveness, reliability, work quality, and the speed of performance of maintenance task. Some of the factors that the technique will consider during its calculations are:

- individual proficiency
- team proficiency
- time constraints
- stress
- use or non-use of protective clothing and oxygen masks
- ambient noise and heat levels
- training
- exposure levels

The purpose of this interview is to determine what information you think is most useful as output from such a technique and what variables you think affect nuclear power plant maintenance most. This information will be used in the development of the technique. The information you provide will be held confidential and will not be identified with you or your organization.

Code # ______ ______ ______

Your name: __________________________ Date: ____________________

(4-5) Months of experience in nuclear power or related fields?

__________ __________

(6) Title of your current position and your work location?

Position title: _______________ Location: ________________

(7) Which of the following most closely describes the organization for which you work?

______ (1) engineering/design organization

______ (2) nuclear power plant

______ (3) regulatory agency

(8) What are the two types of maintenance task that you think are most likely to be done wrong in a NPP? (SHOW CARD)

______ (1) fault diagnosis

______ (2) repair

______ (3) replacement

______ (4) checkout

______ (5) calibration

______ (6) preventive maintenance

______ (7) inspection

______ (8) other; specify _____________________________________

(9) In light of the fact that the assignment of personnel to maintenance tasks is often restricted by radiation and prior exposure levels, do you consider it important that such a technique provide the information required for creating the most cost/efficient team for a maintenance action while keeping the exposure limits within bounds?

_______ (1) yes, such information is important

_______ (2) no, such information is not important

_______ (3) don't know or no response

(10) (If "yes" to #9, answer #10; otherwise skep to #11.)
How useful would that information be? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(11) Suppose that such a technique provided information about how much time should be allowed for a given maintenance task for teams of different training/proficiency. How useful would such information be? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(12) Assume that the results from the technique indicated the best team composition for completing a maintenace task properly and within the desired time limits. How useful would the information be? (SHOW CARD)

______ (1) extremely useful

______ (2) very useful

______ (3) useful

______ (4) slightly useful

______ (5) not useful

(13) Suppose that the technique provided information on how the time allowance should be adjusted if less than the "ideal" team is available. How useful would such information be? (SHOW CARD)

______ (1) extremely useful

______ (2) very useful

______ (3) useful

______ (4) slightly useful

______ (5) not useful

(14) Changes in manning can also affect how well a task is performed. Would it be useful to have information on how different levels of team proficiency may affect the probability of incorrectly performing a task? How useful would this information be? (SHOW CARD)

______ (1) extremely useful

______ (2) very useful

______ (3) useful

______ (4) slightly useful

______ (5) not useful

(15) A trade-off exists between allowing a less than efficient team perform a task and contracting the work out. How useful would information pertinent to this trade-off be? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(16) Errors can occur in the performance of any maintenance task and some of these are not found until later. How useful would you find information of the likelihood of undetected errors? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(17) Of course errors or successes are affected by a great number of factors. Some of these are physical factors, e.g., noise, heat, protective clothing, time constraints, and working in cramped quarters. How useful would information be about how the noise level affects the performance of a given maintenance action? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(18) Assume that the technique provided information about errors/successes as they are affected by the temperature level during task performance. How useful would such information be to you? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(19) Assume that the technique provided information about errors because of the wearing of safety clothing? Would such information be useful?

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(20) Sometimes maintenance tasks must be performed in a tight or cramped place. Would it be useful for you to have information on how these conditions affect errors or successes? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(21) The relationship between errors or successes can also be shown against time constraints. Would this information be useful to you?

_____ (1) extremely useful

_____ (2) very useful

_____ (3) useful

_____ (4) slightly useful

_____ (5) not useful

(22-26) Consider the five physical factors and rank order them by their contribution to the possibility of errors? Assign "1" to the factor which contributes the most, "2" to the next highest contributor, and so on. (SHOW CARD)

Physical factors

_____ (22) noise

_____ (23) heat

_____ (24) safety clothing

_____ (25) time constraints

_____ (26) tight working places

(27-31) Probably all of the information that the technique might make available concerning the relationship between each physical factor and errors or successes would not be equally useful to you. Therefore, rank order the factors to reflect the relative usefulness of the information to you? Assign a "1" to indicate what information you would most want to have, a "2" to the second most useful, and so on.

_____ (27) information relating errors or successes to noise

_____ (28) information relating errrors or successes to heat

_____ (29) information relating errors or successes to safety clothing

_____ (30) information relating errors or successes to time constraints

_____ (31) information relating errors or successes to tight working places

(32) Besides the physical factors, there are any number of situational factors that might influence the performance of maintenance tasks. How useful would information be on how stress might affect performance?

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(33) The technique can also address the issue of how the morale of an assigned maintenance team might affect their performance. Would this information be useful?

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(34) Training is another factor that affects performance. The technique can provide information on how different levels of training and training emphasis may effect errors or successes. How useful would this information be?

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(35) How about information on how the level of proficiency of team members relates to errors/successes? Is information of this type useful?

______ (1) extremely useful

______ (2) very useful

______ (3) useful

______ (4) slightly useful

______ (5) not useful

(36-39) In the case of a choice among the information on the effects of stress, training, morale, or proficiency, which would you find most useful? To show this, rank order the factors assigning a "1" to the factor you find most useful, a "2" to the second most useful and so on. (SHOW CARD)

______ (36) stress

______ (37) training

______ (38) morale

______ (39) proficiency

(40-46) Of course another way to affect performance is through the selection process. This concerns what an individual brings to a work situation with him. Personnel could be selected for any number of reasons, some of which are listed below. Which two are most important? (CHECK TWO) (SHOW CARD)

______ (40) intelligence

______ (41) mechanical ability

______ (42) determination to get things done

______ (43) ambition

______ (44) prior experience

______ (45) precision (accuracy) in performing a task

______ (46) others; specify ______________________________

(47-53) Now, check the two which are most important once a person is actually on the job? (CHECK TWO) (SHOW CARD)

______ (47) intelligence

______ (48) mechanical ability

______ (49) determination to get things done

______ (50) ambition

______ (51) prior experience

______ (52) precision in performing tasks

______ (53) other; specify ______________________________

(54) Assuming you are interested in where maintenance errors are made, would the information concerning errors be more useful if it were given by (1) overall task, (2) the individual actions which constitute the task; i.e., subtasks, or (3) both.

______ (1) by task

______ (2) by subtask

______ (3) both

(55) How about if the subtasks were rank ordered from the most likely to the least likely to be performed incorrectly? How useful would that be? (SHOW CARD)

______ (1) extremely useful

______ (2) very useful

______ (3) useful

______ (4) slightly useful

______ (5) not useful

(56) The time of day can influence the errors or successes in the performance of tasks. An example of a time of day affect might be a reduced alertness in the late night/early morning hours. Information of this type might be relevant to policy matters. Would information be useful on how the time of day affects performance?

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(57) The number of hours worked prior to starting work on a task may affect errors or successes in performance of a current maintenance task, e.g., fatigue, both mental and physical, increases with cumulative hours worked and the number of successive days since the last day off. Is information about how fatigue may affect the performance of a given technician or team of technicians useful?

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(58-62) Of the following factors, which do you think are more likely to lead to an undetected error? Rank order them by their potential contribution assigning "1" to most important, "2" to the score, and so on. (SHOW CARD)

_______ (58) poor system design

_______ (59) use of inappropriate tools

_______ (60) not following procedures

_______ (61) poorly documented procedures

_______ (62) others; specify ________________________________

(63) Risk can be conceived as the degree of danger to the health of the public due to various malfunctions of the system. Many tasks required in the maintenance of nuclear power plants have an inherent risk potential. How useful would information about the risk associated with each task and subtask be? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(64) Maintenance can be conceptualized as preventive maintenance or demand maintenance. Preventive maintenance is work performed to ensure that equipment or system components will continue to function properly. It is usually performed according to a prescribed schedule. Demand maintenance is performed to correct a specific malfunction. What percentage of the maintenance tasks is preventive in nature? (SHOW CARD)

_______ (1) above 90%

_______ (2) 75 to 89%

_______ (3) 50 to 74%

_______ (4) 25 to 49%

_______ (5) less than 24%

(65) What percentage of preventive maintenance tasks are associated with public risk?

_______ (1) above 90%

_______ (2) 75 to 89%

_______ (3) 50 to 74%

_______ (4) 25 to 49%

_______ (5) less than 24%

(66) It is possible to conceive that the same maintenance task, e.g., rebuilding a pump, may have different success or error potentials associated with it when performed as preventive maintenance rather than demand maintenance. How useful would it be to have the error or success information for each maintenance task when performed under a preventive schedule and a demand schedule?

_____ (1) extremely useful

_____ (2) very useful

_____ (3) useful

_____ (4) slightly useful

_____ (5) not useful

(67-70) Rank order the following personnel types as how much an error committed by them would contribute to public risk? Assign "1" to the highest contributor, "2" to the next, and so on.

_____ (67) mechanics

_____ (68) electricians

_____ (69) instrument and control technicians

_____ (70) other; specify _______________

(71) One concern in a NPP is the safety of the maintenance personnel who complete a task. For example, the absorbed radiation is a safety item. How useful would the information be if the technique provided information about the safety hazard to the maintainer?

_____ (1) extremely useful

_____ (2) very useful

_____ (3) useful

_____ (4) slightly useful

_____ (5) not useful

(72) Reliability of a system has to do with how well and how predictable it functions, along with its subsystems, including the human element. No doubt, for each item of equipment, there is some index of reliability. Would a similar reliability index for each maintainer task be useful to you?

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(73) The computer oriented technique for analyzing the performance of maintenance personnel could supply the information in printed form or on a cathode ray tube terminal. Which would you find more convenient?

_______ (1) printout

_______ (2) cathode ray tube

(74) How useful would you find a computer oriented technique which would provide the type of information discussed in aspects of this interview? (SHOW CARD)

_______ (1) extremely useful

_______ (2) very useful

_______ (3) useful

_______ (4) slightly useful

_______ (5) not useful

(75) Is there anything else you would like to say in regard to a technique for analyzing nuclear power plant maintenance personnel performance?

____

____

____

APPENDIX C

MAIL QUESTIONNAIRE

CONDITIONS SURVEY

Applied Psychological Services, under contract with the Nuclear Regulatory Commission and working through the Oak Ridge National Laboratory, is producing a computer-oriented technique for analyzing the performance of nuclear power plant maintenance personnel. This includes mechanical, instrument and control, and electrical personnel. The technique will describe how various personnel, environmental, and equipment design factors contribute to such considerations as cost/effectiveness, reliability, work quality, and the speed of performance of maintenance tasks. Some of the factors that the technique will consider during its calculations are:

- individual proficiency
- team proficiency
- time constraints
- stress
- use of protective gear and respirator
- environmental factors
- training
- task demands

The purpose of this questionnaire is to determine what information you think is most useful as output from such a technique and what variables you think affect nuclear power plant maintenance. This information will be held confidential and will not be identified with you or your organization.

Throughout this questionnaire, whenever the word "technique" is used, it refers to the computer-oriented method for analyzing maintenance performance.

Any comments or clarifications to any of your answers can be entered on the last page which is supplied for the purpose of allowing you to elaborate upon your answer to any question.

Instructions for returning the form are attached to the last page of the form.

(1-3) Code # ___ ___ ___

(4-6) ___ ___ ___ Months of experience in nuclear power.

(7) ____ Your current position title.

(8) Which of the following most closely describes the organization you work for ?

________ (1) engineering/design organization

________ (2) nuclear power plant

________ (3) regulatory agency

(9) What is your primary area of responsibility?

________ (1) mechanical maintenance

________ (2) electrical maintenance

________ (3) instrument and control maintenance

________ (4) quality assurance

________ (5) other; specify ____________________________________

____

The computer-based technique (described on the cover page) for analyzing maintenance performance may prove useful to a wide variety of people concerned with nuclear power. We need to know the kind of information that would be most useful.

This section of the questionnaire is concerned with the nature of the information that would be useful.

(10) How useful would information concerning the composition of the most cost/effective team for assignment to a maintenance action be?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(11) In view of the fact that the assignment of personnel to maintenance tasks is often restricted by radiation and prior exposure levels, how useful would information be about the composition of cost/effective teams while keeping exposure levels within bounds?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(12) Suppose the computer-based technique provided information about how much time is generally required for a given maintenance task for teams of different training/proficiency. How useful would such information be?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(13) The technique could indicate the optimum number of technicians to be included in a team in order to complete a maintenance task properly and within the desired time limits. How useful would the information be?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(14) The computer-based technique could provide specific information on the best team composition, such as the mix of skills and proficiency levels, to perform a task properly and within the desired time limits. How useful would such information be?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(15) Suppose that the technique provided information on how the time allotted for a maintenance task should be adjusted if a team which is less qualified than the "ideal" team is assigned. How useful would such information be?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(16) Changes in manning can affect how well a task is performed. Would it be useful to have information on how different levels of team proficiency may affect the probability of incorrectly performing a task? How useful would this information be?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(17) A trade-off exists between allowing a less than efficient "in house" team perform a task and contracting the work out. How useful would information pertinent to this trade-off be?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(18, 19) Of course another way to affect performance is through the selection process. This concerns what an individual brings to a work situation with him. Personnel could be selected on any number of criteria. Some of these criteria are listed below. Which two are most important? (CHECK TWO)

________ (1) intelligence

________ (2) mechanical ability

________ (3) determination to get things done

________ (4) positive work attitude

________ (5) prior experience

________ (6) precision (accuracy) in performing a task

________ (7) others; specify ______________________________

____

The performance of maintenance tasks is affected by a great number of factors. Some of these are physical factors, e.g., noise, heat, protective gear, time constraints, and working in cramped quarters.

(20) How useful would information be about how noise level affects the performance of a given maintenance action?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(21) Assume that the technique provided information about performance as it is affected by the temperature level during task performance. How useful would such information be?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(22) Assume that the technique provided information about performance as it is affected by the wearing of safety and/or protective gear. Would such information be useful?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(23) Sometimes maintenance tasks must be performed in a tight or cramped place. Would it be useful for you to have information on how these conditions affect performance?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(24) The relationship between performance can also be shown against time constraints. Would this information be useful?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(25-29) Perhaps all of the information that the technique might make available concerning the relationship between each physical factor and performance would not be equally useful to you. Therefore, rank order the factors to reflect the relative usefulness of the information to you. Assign a "1" to indicate what information you would most want to have, a "2" to the second most useful, and so on.

_________ (25) information relating performance to noise

_________ (26) information relating performance to heat

_________ (27) information relating performance to safety gear

_________ (28) information relating performance to tight working places

_________ (29) information relating performance to time constraints

(30) Besides the physical factors, there are any number of situational factors that might influence the performance of maintenance tasks. How useful would information be on how stress might affect performance?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(31) The technique can also address the issue of how the morale of an assigned maintenance team might affect their performance. Would this information be useful?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(32) Training is another factor that affects performance. The technique can provide information on how different levels of training and training emphasis may affect errors or successes. How useful would this information be?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(33) How about information on how the level of proficiency of team members relates to performance? Is information of this type useful?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(34-37) In the case of a choice among the information on the effects of stress, training, morale, or proficiency on performance effectiveness, which would you find most useful? To show this, rank order the factors assigning a "1" to the factor you would find most useful, a "2" to the second most useful and so on.

_________ (34) stress

_________ (35) training

_________ (36) morale

_________ (37) proficiency

(38) The time of day can influence the errors or successes in the performance of tasks. An example of a time of day effect might be a reduced alertness in the late night/early morning hours. Would information be useful on how the time of day affects performance?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

(39) The number of hours worked prior to starting work on a task may affect errors or successes in performance of a current maintenance task, e.g., fatigue, both mental and physical, increases with cumulative hours worked and the number of successive days since the last day off. Is information about how fatigue may affect the performance of a given technician or team of technicians useful?

_________ (1) extremely useful

_________ (2) very useful

_________ (3) useful

_________ (4) slightly useful

_________ (5) not useful

This section centers on human error. The questions are concerned with sources of human error and the type of information you might find useful about human error, risk, and safety.

(40, 41) What are the two types of maintenance task that are most likely to be done wrong in a nuclear power plant? (CHECK TWO)

_______ (1) fault diagnosis

_______ (2) repair

_______ (3) replacement

_______ (4) checkout

_______ (5) calibration

_______ (6) preventive maintenance

_______ (7) inspection

_______ (8) other; specify ______________________________

____

(42-46) Consider the following five physical factors and rank order them by their contribution to the possibility of error. Assign "1" to the factor which contributes the most, "2" to the next highest contributor, and so on.

_______ (42) noise

_______ (43) heat

_______ (44) safety gear

_______ (45) time constraints

_______ (46) tight working places

(47) Assuming that you are interested in where maintenance errors are made, would the information concerning errors be more useful if it were given by: (1) overall task, (2) the individual actions which constitute the task; i. e., subtasks, or (3) both.

________ (1) by task

________ (2) by subtask

________ (3) both

(48) How about if the technique rank ordered subtasks from the most likely to the least likely to be performed incorrectly? How useful would that be?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(49) Errors can occur in the performance of any maintenance task and some of these are not found until later. How useful is information on the likelihood of undetected errors?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(50-55) Of the following factors, which do you think are more likely to lead to an undetected error? Rank order them by their potential contribution assigning "1" to most important, "2" to second most, and so on.

________ (50) poor system design

________ (51) use of inappropriate tools

________ (52) not following precedures

________ (53) poorly documented procedures

________ (54) inadequate supervision and/or supervisory check

________ (55) lack of experience

(56) Risk can be conceived as the degree of danger to the health of the public due to various malfunctions of the system. Some tasks required in the maintenance of nuclear power plants have an inherent risk potential. How useful would information about the risk associated with each task and subtask be?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(57) Maintenance canbe conceptualized as preventive maintenance or demand maintenance. Preventive maintenance is work performed to ensure that equipment or system components will continue to function properly. It is often performed according to a prescribed schedule. Demand maintenance is performed to correct a specific malfunction. What percentage of the maintenance tasks is preventive in nature?

________ (1) 80-100%

________ (2) 60-79%

________ (3) 40-59%

________ (4) 20-39%

________ (5) 0-19%

(58) If performed incorrectly, what percentage of maintenance tasks (both preventive and demand) would lead to an event report, site emergency, or general emergency?

________ (1) 80-100%

________ (2) 60-79%

________ (3) 40-59%

________ (4) 20-39%

________ (5) 0-19%

(59) It is possible to conceive that the same maintenance task, e.g., rebuilding a pump, may have different success or error potentials associated with it when performed as preventive maintenance rather than demand maintenance. How useful would it be to have the error or success information for each maintenance task when performed under a preventive schedule and a demand schedule?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(60-63) Rank order the following personnel types by how much an error committed by them would contribute to public risk. Assign "1" to the highest contributor, "2" to the next, and so on.

________ (60) mechanical technicians

________ (61) electrical technicians

________ (62) instrument and control technicians

________ (63) quality assurance inspectors

(64) One concern in a nuclear power plant is the safety of the maintenance personnel who complete a task. For example, nuclear plants contain radiation, chemical, and steam hazards. How useful would it be if the technique provided information about the safety hazard to the maintainer?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(65) Reliability of a system has to do with how well and how predictably it functions, along with its subsystems. No doubt, for each item of equipment there is some index of reliability. Would a similar reliability index for each mainenance task be useful?

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

(66) The computer-oriented technique for analyzing the performance of maintenance personnel could supply the information in printed form or on a cathode ray tube terminal. Which would you find more convenient?

________ (1) printout

________ (2) cathode ray tube

________ (3) both

(67) How useful would you find a computer-oriented technique which would provide the type of information discussed in aspects of this questionnaire? Answer this question in relation to your own particular sphere of interest and not as it may generally apply to the nuclear power generating industry.

________ (1) extremely useful

________ (2) very useful

________ (3) useful

________ (4) slightly useful

________ (5) not useful

Enter anything else you would like to say about any question by question number. Also, enter anything else you would like to say in regard to a technique for analyzing nuclear power plant maintenance personnel performance.

COMMENTS

Thank you for your cooperation.

APPENDIX D

KEY WORDS AND QUESTION NUMBERS

Key Word	Question Number
Cost-effective team	10
Exposure	11
Time required	12
Number	13
Team composition	14
Time adjustment	15
Probability of error	16
Contracting out	17
Noise	20
Temperature	21
Gear	22
Tight places	23
Time constraints	24
Stress	30
Morale	31
Training	32
Proficiency	33
Time of day	38
Fatigue	39
Subtask success probability	48
Undetected errors	49
Risk	56
Demand vs. PM	59
Safety	64
Human reliability	65
Overall	67

APPENDIX E

DETAILS OF THE MAPPS PROGRAM PLAN

E.1. General Program Plan

The program plan for the MAPPS model spans a 38-month period of time and consists of three distinct phases; the development phase, the validation phase and the dissemination phase. Model development will be initiated in May 1982 and dissemination will be completed during June 1985. Two releases of the model are scheduled to occur during this time period. The first will be a debugged/sensitivity tested/calibrated version of the model (version 1) and will occur at the end of the development phase (January, 1984). The second will be a validated version of the model and will occur at the end of the validation phase (January, 1985).

The program will also be subject to peer review by a member of subject matter experts. Several peer review group meetings are tentatively scheduled as indicated in Figure E.l. Each of the phases will now be discussed in detail.

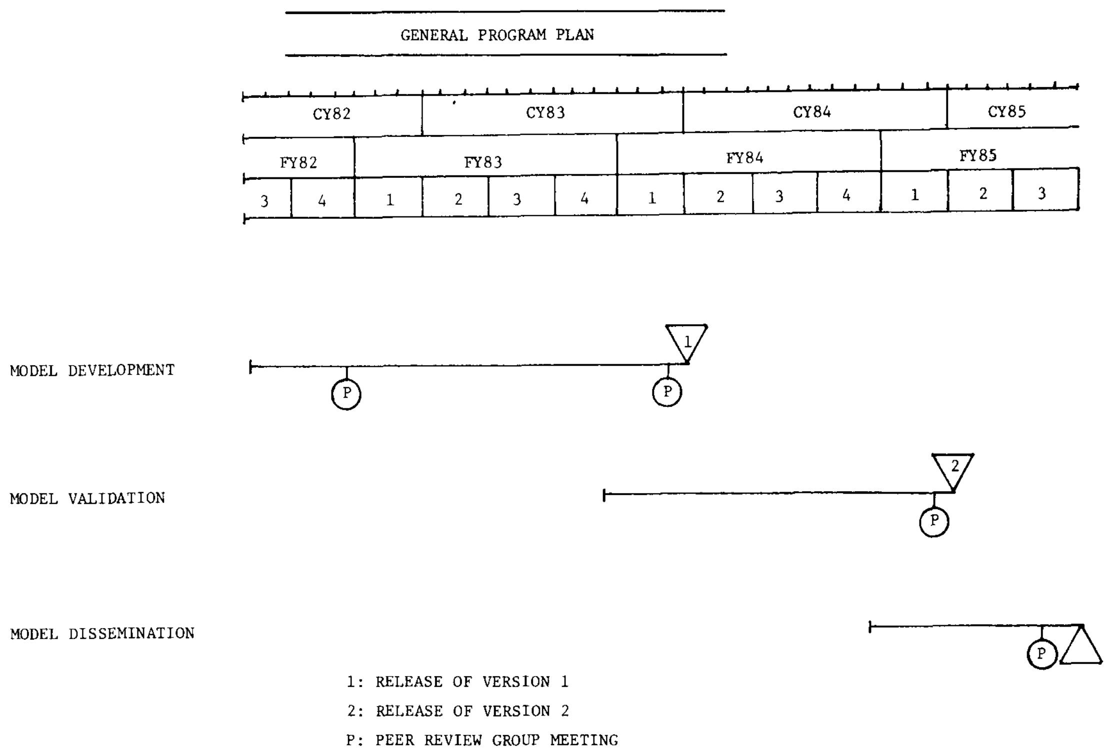

Fig. E.1. General Program Plan for MAPPS.

E.2. Model Development Phase

The model development phase of the program will span a 20 month period of time and will culminate with the release of version 1 of the model (debugged/sensitivity tested/calibrated version). The objective of this phase of the program is to formulate and develop a computer simulation model to predict the reliability of nuclear power plant maintenance personnel in the performance of maintenance tasks. Efforts during this phase will be concerned primarily with the completion of a detailed, sequenced logic for, programming of, and sensitivity testing of the model.

The formulation of the general structure of the model will be accomplished by identifying various applicable human behavioral theories, reliability theories and existing human behavioral methodologies through the completion of appropriate literature analyses and applying them to the nuclear power plant maintenance context. Appropriate theories (e.g., psychosocial, maintenance, reliability, probabilistic) will be applied in developing the mathematical and digital logic for each of the model's internal and output variables. Explicit functional relationships between model variables will be developed and will be verified both independently and within the context of the model. Decisions during the formulation stage of model development will be guided by criteria such as reasonableness of output over all model parameter ranges, processing economy, indifference to trivial effects, incorporation of accepted pschosocial theory, and the importance of output results.

As the general structure and logic of the model becomes more complete, model development will be initiated. This effort will include the choice of programming language, object computer and type of terminal, as well as determination of the detailed program structure, assessment of data element names, definitions of context, organization and formatting of input and output data. The program will be organized into discrete subroutines as modules which possess logical and processing coherency. Structured programming techniques will be utilized in the development of each of the model's modules and a detailed definition of each module will be prepared. To facilitate program checking, a queue-trail list of the module numbers in the sequence called and a count of the number of times each is entered will be maintained. Test data suitable for program checking will be prepared and run against the model for model program debugging.

Once debugging of the model is completed, sensitivity testing will be initiated. Sensitivity testing will be done in two parts. The first part will include an extensive set of computer runs for a selected sample of tasks in which each parameter in turn is subjected to assignment of values over its entire range of admissable values in a systematic variation. These will be analyzed and compared against a baseline simulation of nominal parameter values. The first part of sensitivity testing will be done on the model as a whole and for each module and subroutine of the model. In order to evaluate the treatment of parameter interdependencies. The second part of sensitivity testing will include running the model against actual task analyses data gathered in the field. The tasks selected will meet such general criteria as: (1) being common to both pressurized water reactor plants and boiling water reactor plants, (2) performed by two or more persons, (3) utilizes more than one technician job speciality and (4) performance time is greater than two hours. Other criteria will also be applied in order to select a set of tasks which will tax the limits of the model. Once the task analyses data has been gathered, applied to the model and results generated, they will be composed to a baseline of expected results in order to determine the reasonableness of the model output. The baseline of expected results will be a limited set of actual performance data gathered during task analyses. A more formal and indepth comparison to actual performance data will be done during the validation phase of this program.

A limited amount of modification to the model is expected during sensitivity testing. For example, rescaling of some variables or the dampening of various effects within the model may be required. Following the incorporation of any modifications to the model, additional runs will be made and again compared to the baseline of expected results. These comparisons will be done in order to provide any final calibration to the model prior to the validation effort. Once the model has been calibrated, it will be considered ready for validation and the state of the model at this time will be "frozen" as is, for validation.

The formulation, development, sensitivity testing and calibration of the model will be fully documented. In addition, the structure and logic of the model will be delineated along with all functional relationships between the various modules within the model. Input and output parameter will be specified as well as a description of the various types of model output formats. Model limitations will also be discussed and detailed instructions for model usage will be provided in conjunction with several illustrative examples of its application.

The release of version 1 of the model (debugged, sensitivity tested and calibrated) will be at the end of the development phase of this program (January, 1984). Its release will be limited to the NRC, with a more general release of the model planned following its validation. A detailed plan for the development phase of the program including subtasks is given in Figure E.2.

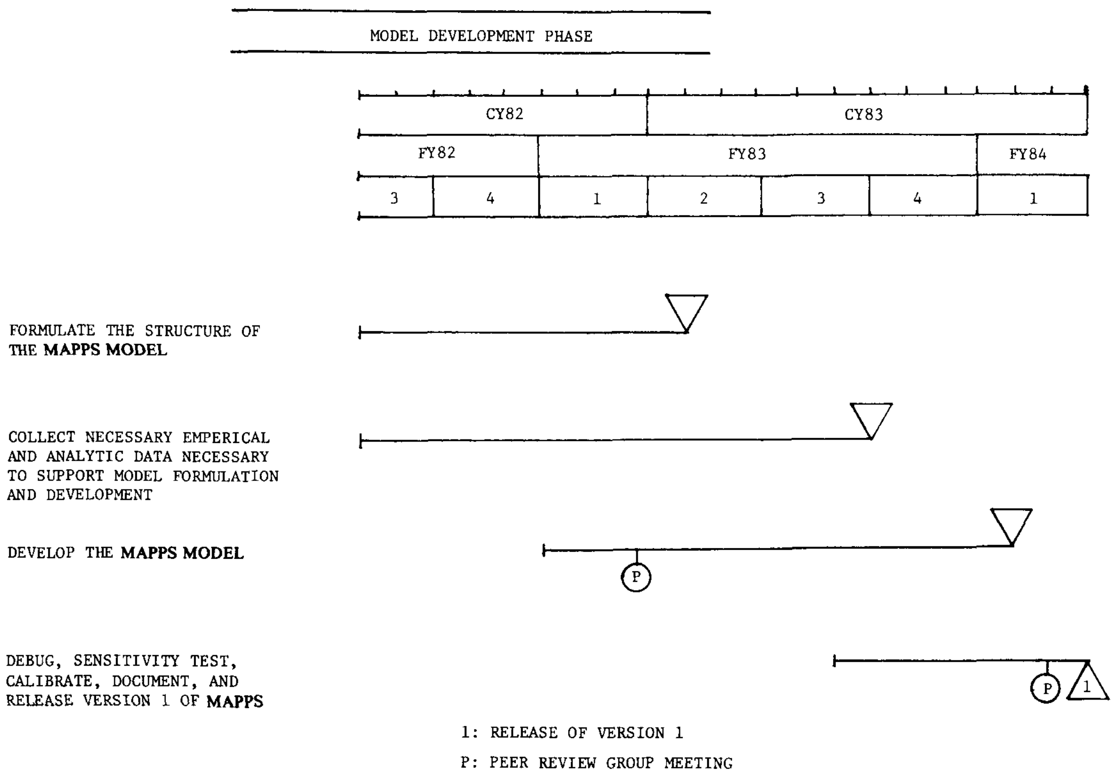

Fig. E.2. Plan for Model Development.

E.3. Model Validation Phase

The model validation phase of the program will span a 16 month period of time and will culminate with the release of version 2 of the model (validated version). The objective of this phase of the program is to demonstrate that the behavioral inferences produced by the model are in agreement with rated/observed data and information pertaining to the behavior of maintenance personnel performing maintenance tasks at a nuclear power plant.

Validation of the MAPPS model will be done according to a validation plan developed early in this phase of work. The validation plan developed will reflect such practical criteria as cost, availability of personnel at nuclear facilities, availability of subject matter experts, time considerations, etc. In addition, the plan will be developed in such a way as to maximize the model's credibility with respect to the NRC and the nuclear industry in general. This acceptance criteria will be accomplished by incorporating, as possible, important probabilistic risk assessment concerns and practical utility concerns within the validation effort. The validation plan will also be mindful of demonstrating the model's usefulness. The results of the model will be evaluated statistically for significant correlations with rated/observed data. In addition, relevant metrics for probabilistic risk assessment and human reliability analyses will be indicated.

The final validation plan will consist of two primary sections. The first will be a comprehensive validation strategy which will be the end result of the consideration of a number of validation approaches. The second section will be a specification of the tasks to be addressed during the validation effort. The final validation approach and validation tasks that are chosen will reflect the criteria of practicality, acceptance and usefulness discussed earlier. A tentative list of validation approaches is presented in Table E.1.

Table E.1. Tentative Validation Approaches

1. Validation against actual field data collected by the project team.
2. Validation against information obtained from subject matter experts using a Delphi Technique.
3. Validation against data and information obtained from existing data sources such as data bases, maintenance records, etc.
4. Validation against information obtained from interviews (task-talk through) with plant maintenance personnel.
5. Validation against information obtained from maintenance simulators, or walk-throughs on system mockups.

The validation strategy will provide a detailed description of one or more of the validation methods chosen for this program. It will also outline means of implementing each of these methods and will provide a summary list of information that is expected to result from the validation strategy. Each of the validation methods implemented will result in two primary sets of information. First, a comprehensive profile of maintenance performance with respect to the validation tasks will be produced. Secondly, complete set of team variables for input into the MAPPS model will be generated.

Task analyses for the validation tasks will be completed in order to supply the needed task-specific data for input into MAPPS. Simulation of the validation tasks will be accomplished, and will generate a simulation profile of maintenance performance. A statistical comparison between simulation results to the results obtained via the validation approaches will be made to determine if they are statistically different. The comparison of the simulated maintenance profile to that generated by one validation approaches will be a measure of the models external validity, i.e., now model results compare to actual results. The models internal validity will also be examined. This will be accomplished by examining the module to model output relationships for significant and meaningful correllations as well as examining the importance relationships between modules.

The validation phase of this program will be fully documented. The NUREG/CR to be published for the validation phase will provide detailed information concerning the implementation of the validation plan, provide statistical results concerning the external and internal validity analyses and identify area requiring improvements, if any. Two other publications will also be issued. The first of these will be a detailed description of the model. This report will detail the model logic including all subroutine and functional relationships used within the model. Model option as well as model input parameters will be summarized. All output option will be discussed and an illustration of each, with interpretation, will be provided. The second report will be a users manual for the model. Detailed illustrations for running the model will be provided which will include a discussion of required input data, default values, model parameters, and output data. The user manual will include illustrative examples with a discussion of their output. One or more fully set up model runs will also be provided.

The fully validated model will be released to the NRC (January, 1985) and the public. The model will be accompanied by a library consisting of task analysis data and maintenance team data for up to five maintenance tasks. The purpose of this library is to provide the new user with a complete set of information and data that may be used for instructional purposes.

A peer review meeting will be held shortly before the release of version-2. The validation effort will be reviewed and evaluated by the group. It will also be the purpose of this meeting to identify areas that may require improvement in the model and to suggest future areas of research which may benefit from application of the model. A detailed plan for the validation phase of the program including subtasks is given in Figure E.3.

E.4. Model Dissemination Phase

The model dissemination phase of the program will span a ten month period of time. The objective of this phase of the program is to transfer to the NRC and to other potential users the methodology developed within this program and to assure its effective implementation by users.

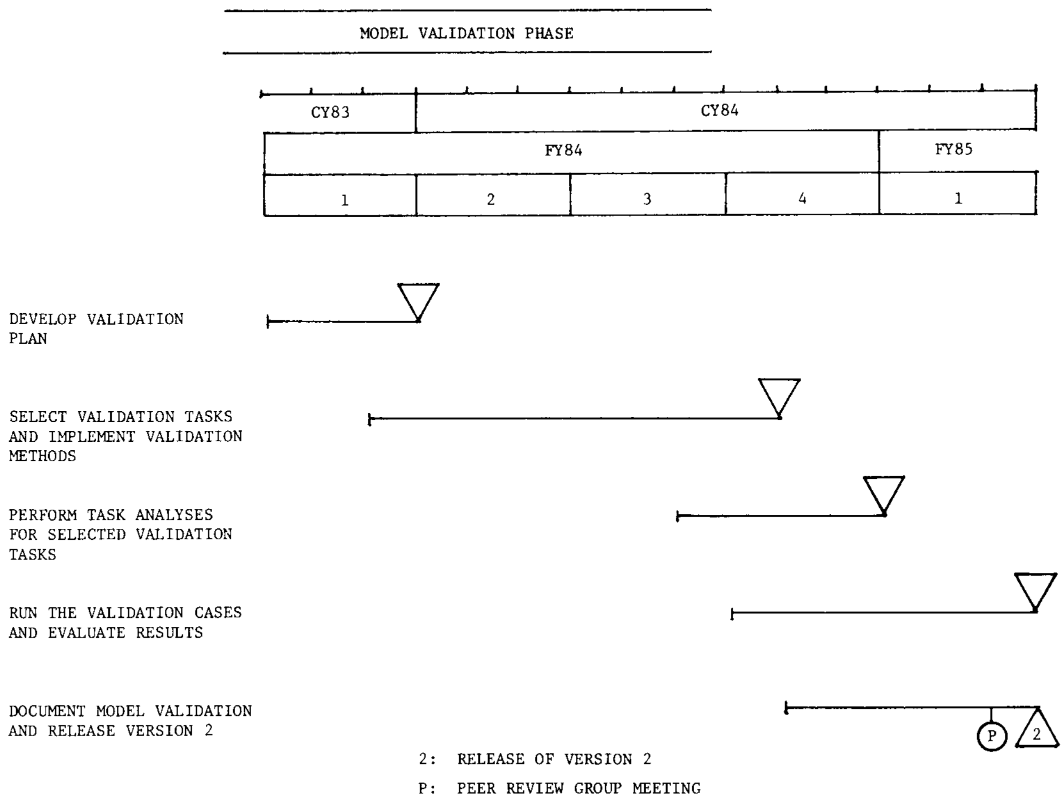

Fig. E.3. Plan for Model Validation.

The transfer of the methodology will be done according to a dissemination plan developed early in this phase of the program. The plan will outline a comprehensive workshop that will have as its goal the transfer of the model to user and its efficient implementation by users. The workshop will be conducted by members of the project staff who have been intimately involved in the development of the model. Every effort will be made to assure that participants of the workshop will acquire a detailed working knowledge of the model, requirements for running of the model, and an ability to interpret properly the metrics produced by the model. The workshop will allow the user to gain hands-on experience at running the model. Demonstrations of model usage for various model modes such as the generation of reliability data, sensitivity analyses, identification of problem areas within maintenance task, etc. will be provided. In addition, a portion of the workshop will be dedicated toward performing successful task analyses for the generation of input data required by the model.

Each participant in the workshop will receive a MAPPS users manual, model description report and a set of workshop materials including class notes, completed example problems, illustration of task analyses, examples of model outputs, example of output interpretation, and examples to be worked in class.

A peer review group meeting will be held in conjunction and subsequent to the workshop to evaluate the methodology transfer, and to suggest areas of improvement if any. The plan for the dissemination phase is presented in Figure E.4.

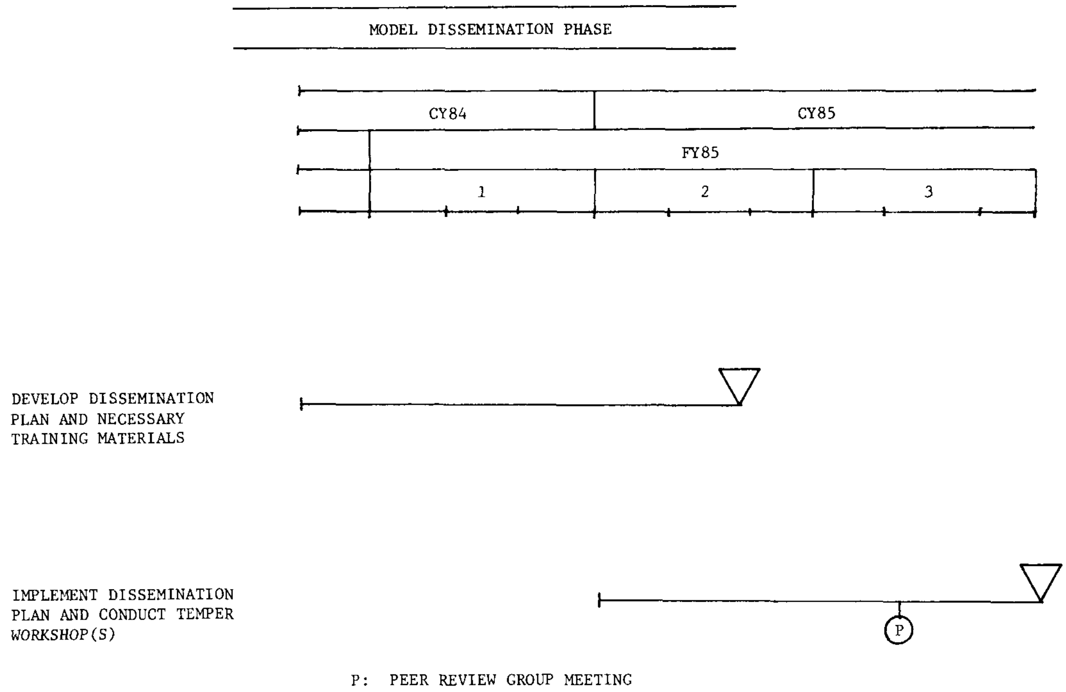

Fig. E.4. Plan for Model Dissemination.

E.5. Summary of the Program Plan

Appendix E of this report was provided a detailed program plan for the development of the MAPPS model. The total program time plan is 38 months and will result in a fully validated version of the model capable of supplying important human reliability metrics to the NRC and to other organizations needing human reliability data for the maintenance context. A detailed program plan schedule by task and subtask is presented in Figure E.5.

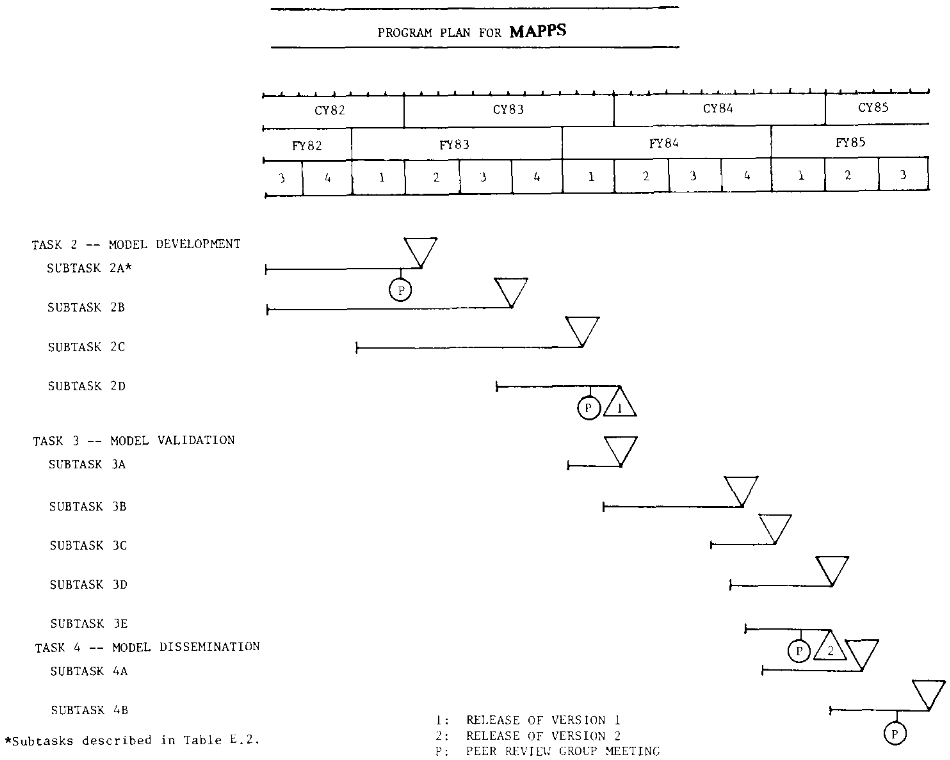

Fig. E.5. Detailed Program Plan for MAPPS.

Table E.2. Subtask Descriptions

Subtask-2A	Formulate the structure of the MAPPS model.
Subtask-2B	Collect necessary empirical and analytic data to support model formulation and development.
Subtask-2C	Develop the MAPPS model.
Subtask-2D	Debug, sensitivity test, calibrate, document and release version 1 of MAPPS.
Subtask-3A	Develop validation plan.
Subtask-3B	Select validation tasks and implement validation methods.
Subtask-3C	Perform task analyses for elected validation tasks.
Subtask-3D	Run validation cases and evaluate results.
Subtask-3E	Document model validation and release version 2.
Subtask-4A	Develop dissemination plan and necessary training materials.
Subtask-4B	Implement dissemination plan and conduct MAPPS workshop(s).

NUREG/CR-2669
ORNL/TM-8300
Distribution Category RX

INTERNAL DISTRIBUTION

1–2. L. S. Abbott
3. W. B. Cottrell
4. R. M. Davis
5. G. F. Flanagan
6. D. N. Fry
7. L. H. Gray
8–12. P. M. Haas
13. R. A. Kisner
14–18. H. E. Knee
19. A. L. Lotts
20. F. C. Maienschein
21. J. J. Manning
22. D. L. Selby
23. W. H. Sides
24. D. B. Trauger
25. A. Zucker
26. P. W. Dickson, Jr. (Consultant)
27. H. J. C. Kouts (Consultant)
28. W. B. Lowenstein (Consultant)
29. R. Wilson (Consultant)
30. Central Research Library
31. Y-12 Document Ref. Section
32. Laboratory Records ORNL, RC
33. ORNL Patent Office

EXTERNAL DISTRIBUTION

34. Office of Assistant Manager for Energy Research & Development, DOE-ORO, Oak Ridge, TN 37830.

35. Director, Division of Reactor Research & Technology, U.S. DOE, Washington, DC 20545.

36. H. L. Parris, Electric Power Research Institute, 3412 Hillview Ave., P.O. Box 10412, Palo Alto, CA 94303.

37. E. P. Wilkinson, Institute of Nuclear Power Operations, 1820 Water Place, Atlanta, GA 30339.

Applied Psychological Services, Inc., Science Ctr., Wayne, PA 19087:

38. William D. Bartter
39. Arthur I. Siegel
40. J. J. Wolf

U.S. Nuclear Regulatory Commission, Washington, D.C. 20555:

Division of Facility Operations:

41. J. P. Jenkins
42. T. G. Ryan
43. J. A. Norberg
44. C. M. Overbey

Division of Human Factors Safety:

45. D. I. Serig
46. H. R. Booher
47. J. J. Kramer
48. D. Jones
49. D. H. Beckham

Division of Risk Analysis:

50. K. Murphy
51. P. W. Baranowsky

52–53. Technical Information Center

54–453. Given distribution as shown in NRC Category RX, Human Factors Safety (NTIS - 10).

NRC FORM 335 (11-81)

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

1. REPORT NUMBER (Assigned by DDC): NUREG/CR-2669 ORNL/TM-8300

2. (Leave blank)

3. RECIPIENT'S ACCESSION NO.

4. TITLE AND SUBTITLE (Add Volume No., if appropriate)
Front-End Analysis for the Nuclear Power Plant Maintenance Personnel Reliability Model

5. DATE REPORT COMPLETED
MONTH: March | YEAR: 1983

7. AUTHOR(S)
Arthur I. Siegel, William D. Bartter, J. Jay Wolf, Helmut E. Knee, Paul M. Haas

DATE REPORT ISSUED
MONTH: June | YEAR: 1983

9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Oak Ridge National Laboratory
P.O. Box X
Oak Ridge, TN 37830

6. (Leave blank)

8. (Leave blank)

12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Nuclear Regulatory Commission
Washington, DC

10. PROJECT/TASK/WORK UNIT NO.
40X-70527

11. FIN NO.
B0461

13. TYPE OF REPORT
NUREG/CR

PERIOD COVERED (Inclusive dates)

15. SUPPLEMENTARY NOTES

14. (Leave blank)

16. ABSTRACT (200 words or less)

The front-end analysis performed for the nuclear power plant maintenance personnel reliability modeling program consisted of three primary tasks which are addressed within this report. The first of these was a front-end user survey which investigated the need for and potential content of a structured methodology for nuclear power plant maintenance. The second task was a literature review of existing human behavioral methodologies and an assessment of their applicability for this program. The third task was the development of a comprehensive program plan for the maintenance reliability model.

Results of these tasks indicated that a computerized model would be very useful and that the type of methodology to be developed should be of the simulation type. This report also provides a 35-month program plan for its development.

17. KEY WORDS AND DOCUMENT ANALYSIS

Nuclear power plant maintenance
Maintenance
Simulation models
Human reliability
Modeling

17a. DESCRIPTORS

17b. IDENTIFIERS/OPEN-ENDED TERMS

18. AVAILABILITY STATEMENT
Unlinited

19. SECURITY CLASS (This report)
Unclassified

20. SECURITY CLASS (This page)
Unclassified

21. NO. OF PAGES

22. PRICE
$

NRC FORM 335 (11-81)