NUREG/CR-4016
BNL-NUREG-51828
Vol. II

APPLICATION OF SLIM-MAUD: A TEST OF AN INTERACTIVE COMPUTER-BASED METHOD FOR ORGANIZING EXPERT ASSESSMENT OF HUMAN PERFORMANCE AND RELIABILITY

VOLUME II: APPENDICES

C.M. Spettell, E.A. Rosa, P.C. Humphreys, and D.E. Embrey

Date Published — October 1986

DEPARTMENT OF NUCLEAR ENERGY, BROOKHAVEN NATIONAL LABORATORY
UPTON, LONG ISLAND, NEW YORK 11973

Prepared for
United States Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Contract No. DE-AC02-76CH00016

NUREG/CR-4016
BNL-NUREG-51828
Vol. II
RX, AN, RG, 1S

APPLICATION OF SLIM-MAUD: A TEST OF AN INTERACTIVE COMPUTER-BASED METHOD FOR ORGANIZING EXPERT ASSESSMENT OF HUMAN PERFORMANCE AND RELIABILITY

VOLUME II: APPENDICES

C.M. Spettell,* E.A. Rosa,* P.C. Humphreys,+ and D.E. Embrey†

+ Department of Social Psychology
London School of Economics and Political Science
Houton Street
London WC2A, England

† Human Reliability Associates, Ltd.
1 School House, Higher Lane, Dalton
Parbold, Lancashire WN8 7RP, England

Manuscript Completed — May 1985
Manuscript Revised — August 1986

Prepared under Contract to
* DEPARTMENT OF NUCLEAR ENERGY
BROOKHAVEN NATIONAL LABORATORY, ASSOCIATED UNIVERSITIES INC.
UPTON, LONG ISLAND, NEW YORK 11973

Prepared for
UNITED STATES NUCLEAR REGULATORY COMMISSION
HUMAN FACTORS AND SAFEGUARDS BRANCH
OFFICE OF NUCLEAR REGULATORY RESEARCH
CONTRACT NO. DE-AC02-76CH00016
FIN A-3272

Available from
Superintendent of Documents
U.S. Government Printing Office
P.O. Box 37082
Washington, DC 20013-7982
and
National Technical Information Service
Springfield, Virginia 22161

ABSTRACT

The U.S. Nuclear Regulatory Commission (NRC) has been conducting a multi-year research program to investigate different methods for using expert judgments to estimate human error probabilities (HEPs) in nuclear power plants. One of the methods investigated, derived from multi-attribute utility theory, is the Success Likelihood Index Methodology implemented through Multi-Attribute Utility Decomposition (SLIM-MAUD). This report describes a systematic test application of the SLIM-MAUD methodology. The test application is evaluated on the basis of three criteria: practicality, acceptability, and usefulness.

Volume I of this report presents an overview of SLIM-MAUD, describes the procedures followed in the test application, and provides a summary of the results obtained.

Volume II consists of technical appendices to support in detail the materials contained in Volume I, and the users' package of explicit procedures to be followed in implementing SLIM-MAUD.

The results obtained in the test application provide support for the application of SLIM-MAUD to a wide variety of applications requiring estimates of human errors.

Previous Reports in the SLIM-MAUD Research Program:

"The Use of Performance Shaping Factors and Quantified Expert Judgments in the Evaluation of Human Reliability: An Initial Appraisal," NUREG/CR-2986, 1983.

"SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgment, Volume I: Overview of SLIM-MAUD and Volume II: Detailed Analysis of the Technical Issues," NUREG/CR-3518, BNL-NUREG-51716, 1984.

CONTENTS

CONTENTS (CONTINUED)

CONTENTS (CONTINUED)

CONTENTS (CONTINUED)

TABLES

FIGURES

ACKNOWLEDGMENTS

A number of individuals and organizations* provided assistance that was essential to the completion of the work presented in this report and we would like to acknowledge their contributions. William J. Luckas, Jr. of Brookhaven National Laboratory (BNL), contract monitor and Dr. Thomas G. Ryan of the U.S. Nuclear Regulatory Commission (NRC), program manager, provided encouragement, support, and guidance. Technical review was provided by Dr. John N. O'Brien of BNL. Clerical assistance was provided by Sharon M. Moore.

Special thanks are due to the eight individuals in the Division of Nuclear Energy, BNL, the two General Physics (GP) instructors, and the six staff members of the NRC who, by participating in the SLIM-MAUD sessions, provided the data analyzed in this report. Recognition is also due to K. Rea of Human Reliability Associates for performing some of the reported analyses, L. Weston of Sandia National Laboratories (SNL) for technical guidance on the task descriptions, and M. K. Comer of GP for arranging for the participation of the GP instructors.

*Work was carried out under the auspices of the U.S. Nuclear Regulatory Commission.

ABBREVIATIONS

ADS	Automatic Depressurization System
ATWS	Anticipated Transient Without Scram
BNL	Brookhaven National Laboratory
BWR	Boiling Water Reactors
DOE	U.S. Department of Energy
HEP	Human Error Probability
HPCI	High Pressure Coolant Injection
HRA	Human Reliability Analysis
LOCA	Loss of Coolant Accident
LOSP	Loss of Off-Site Power
NRC	U.S. Nuclear Regulatory Commission
PRA	Probabilistic Risk Assessment
PSF	Performance Shaping Factor
PWR	Pressurized Water Reactor
RAM	Random Access Memory
RCIC	Reactor Core Isolating Cooling
RHRSW	Residual Heat Removal Service Water
SLC	Standby Liquid Control
SLI	Success Likelihood Index
SLIM-MAUD	Success Index Methodology - Multi-Attribute Utility Decomposition
SNL	Sandia National Laboratories
STAHR	Socio-Technical Approach to Human Reliability

EXECUTIVE SUMMARY

In order to determine the impact of human reliability on nuclear power plant safety, realistic estimates of human error probabilities (HEPs) are needed for probabilistic risk assessments (PRAs). Frequency data for use in estimating HEPs are generally unavailable, or if available, apply to a very limited range of fairly simple actions. To overcome this dilemma, the U.S. Nuclear Regulatory Commission (NRC) has embarked upon a program of research devoted to obtaining estimates of HEPs indirectly, that is, by using expert judgments to arrive at error estimates.

Brookhaven National Laboratory (BNL) developed and evaluated one method of obtaining human reliability estimates from expert judges--the Success Likelihood Index Methodology (SLIM). SLIM comprises a set of procedures based on Multi-Attribute Utility Theory for eliciting and organizing estimates by experts of the probability of success or failure of specific human actions in nuclear power plants.

The feasibility and implementability of SLIM were evaluated in a multiphase investigation. In the first phase, the basic characteristics of SLIM were defined (Embrey, 1983). Phases 2 and 3 consisted of an experimental evaluation and field test of SLIM. In Phase 4, SLIM was linked to an interactive computer program based upon Multi-Attribute Utility Decomposition (MAUD), and procedures for applying the resultant SLIM-MAUD methodology were developed (Embrey, Humphreys, Rosa, Kirwan, and Rea, 1984a; 1984b).

Phase 5, the final phase of the SLIM-MAUD research program, is reported in this two-volume document. Phase 5 was devoted to a systematic test application of the SLIM-MAUD methodology in order to evaluate its practicality, acceptability, and usefulness and to refine the procedures for implementing it.

The practicality of the SLIM-MAUD methodology was evaluated in terms of the costs of implementation, hardware and software requirements, personnel and time requirements, the expandability, transportability and ease of implementation of the methodology, and its ability to interface with the Human Reliability Data Bank. The acceptability of SLIM-MAUD to the scientific community, experts participating in the SLIM-MAUD test application, potential users, NRC and nuclear facilities was evaluated. The usefulness of the SLIM-MAUD methodology was evaluated in terms of the estimates' reliability, face validity and convergent validity.

The test application of SLIM-MAUD was divided into the following stages:

Stage 1 - Selection of Tasks for Assessment in the Test

Thirty tasks were selected for assessment using the SLIM-MAUD methodology. The tasks selected were identical to those employed in Comer, Seaver, Stillwell, and Gaddy's (1984) evaluation of psychological scaling as a method of estimating HEPs for nuclear power plant tasks. Fifteen tasks were designated as Level A and combined BWR plant systems with human actions which represented control room operator duties. Fifteen tasks were designated as Level B and combined equipment components with human actions which represented control room or equipment operator task elements.

<u>Stage 2 - Classification of Tasks into Subsets for Simultaneous Assessment Within SLIM-MAUD</u>

A requirement of the SLIM-MAUD methodology is that tasks be sorted into subsets of 4 to 10 tasks which are reasonably homogeneous with respect to the performance shaping factors (PSFs) presumed to affect task outcome. Therefore, each set of Level A and B tasks were sorted into subsets by a group of experts composed of individuals with PRA expertise, human factors expertise and nuclear power plant operations experience.

<u>Stage 3 - Selection of the Members of the Four Subject Mater Expert Groups for Stage 4</u>

Four groups of subject matter experts were formed composed of individuals with human factors, PRA, or plant operations experience.

<u>Stage 4 - Use of SLIM-MAUD by Each Subject Matter Expert Group for Each Subset of Tasks, Followed by Direct Numerical Assessment of all Tasks in all Subsets by Each Group Member</u>

Each group of experts implemented the SLIM-MAUD methodology with the aid of a facilitator for each subset of tasks resulting in six SLIM-MAUD sessions per group. Then, in order to evaluate the relative usefulness of the SLIM-MAUD methodology as a technique for estimating HEPs, each expert participant used the psychological scaling techniques employed by Comer et al. (1984) to make direct estimates of HEPs for the 15 Level A and 15 Level B tasks. Finally, each expert participant completed an questionnaire to evaluate the SLIM-MAUD methodology in terms of ease of use, ability of the methodology to elicit and organize the judgments of a group of experts and the meaningfulness of the results produced.

<u>Stage 5 - Analysis and Interpretation of Results from SLIM-MAUD Sessions With Respect to the Issues of Practicality, Acceptability and Usefulness</u>

The issues of practicality of the SLIM-MAUD methodology were addressed in qualitative fashion. Formal and informal analyses were carried out to evaluate the acceptability and face validity of the methodology. Correlational and nonmetric multidimensional scaling analyses were conducted to assess the inter-judge reliability of SLIM-MAUD results, and its convergent validity with other subjective techniques for estimating HEPs. Additional analyses were also performed to investigate potential sources of bias in the SLIM-MAUD methodology.

The principal conclusions of this study were:

1. The practicality of SLIM-MAUD was demonstrated with respect to implementation costs, hardware, software, personnel, and time requirements and transportability of the methodology. The expandability of the SLIM-MAUD methodology was supported by the successful assessment of both complex (Level A) and simple (Level B) tasks. SLIM-MAUD can be implemented by a group of subject matter experts after receiving a minimal amount of training.

2. The estimates produced by SLIM-MAUD attained acceptable levels of reliability and showed greater inter-judge consistency than direct estimates of HEPs using psychological scaling techniques.

   To ensure the reliability of SLIM-MAUD results, the following recommendations are made:

   - Tasks to be assessed should be defined as concretely and completely as possible

   - For generic applications, expert groups should first identify a specific plant to typify the range of plants to which the results will be generalized;

   - The expert group should consist of four members. For the assessment of complex tasks, individuals with plant operations experience should form a majority in the group; for simple tasks, individuals with human factors expertise should form a majority in the group.

3. Considerable support was found for the face validity and convergent validity of SLIM-MAUD.

4. Expert participants in the SLIM-MAUD sessions found the SLIM-MAUD methodology easy to use and understand, useful in eliciting and organizing their judgments, and able to produce meaningful results.

5. The SLIM-MAUD methodology enables users to identify which PSFs have the most effect on the SLIs produced, thereby supporting the methodology's acceptability to safety study applications. The HEPs produced by SLIM-MAUD can be used in PRA and in the Human Reliability Data Bank.

Overall, SLIM-MAUD met or exceeded each of the criteria of practicality, acceptability and usefulness. Therefore, it is recommended as a methodology for producing HEPs needed for PRA and for entry into the Human Reliability Data Bank.

1.0 INTRODUCTION

In recent years there has been a widespread and growing concern over human reliability in nuclear power plants, with special concern for obtaining realistic estimates of human error probabilities (HEPs) for probabilistic risk assessments (PRAs). Frequency data from which to estimate human error rates (HERs) is generally unavailable, or if available, pertain to a very limited range of fairly simple actions. To overcome this dilemna, the U.S. Nuclear Regulatory Commission (NRC) embarked upon a program of research devoted to obtaining estimates of human errors indirectly, that is, through techniques using expert judgments to arrive at error estimates. The goal of this research effort is to produce HEP estimates in support of Human Reliability Analysis (HRA) segments of PRAs.

The research program conducted by Brookhaven National Laboratory (BNL) investigated one method of obtaining human reliability estimates from expert judges--the Success Likelihood Index Methodology (SLIM). SLIM comprises a set of procedures based on multi-attribute utility theory for eliciting and organizing experts' estimates of the probability of success or failure of specific human actions in nuclear power plants.

The feasibility and implementability of SLIM were evaluated through a multi-phase investigation. In the first phase, the basic characteristics of SLIM were defined (Embrey, 1983). Phases two and three consisted of an experimental evaluation and field test of SLIM. In phase four, SLIM was linked to a computer-based elicitation procedure based upon Multi-Attribute Utility Decomposition (MAUD) and procedures for applying the resultant SLIM-MAUD methodology were developed. A detailed discussion of phases two through four is reported in Embrey, Humphreys, Rosa, Kirwan, and Rea (1984a; 1984b).

The final phase of the SLIM-MAUD research program, reported in this two-volume document, was devoted to a systematic test application of the SLIM-MAUD methodology and the refinement of the procedures for implementing it. The test application consisted of an evaluation of the practicality, acceptability, and usefulness of the SLIM-MAUD approach for generating estimates of human error probabilities from expert judges.

This is the second volume in the report of the test application of SLIM-MAUD. Volume I (NUREG/CR-4016) presents a general overview of methodologies for systematizing the subjective judgments of experts and reviews the methods and results of the SLIM-MAUD test evaluation.

Volume II contains three appendices. Appendix A: User Manual for Implementing SLIM-MAUD, lists the personnel and equipment needed to conduct a SLIM-MAUD session and contains detailed instructions for using the MAUD programs. Appendix B: Detailed Methods and Results of the Test Application of SLIM-MAUD, presents a detailed description of the methods used in the test application of SLIM-MAUD and the results of the analyses performed to evaluate the practicality, acceptability, and usefulness of SLIM-MAUD. Appendix C:

(Success Likelihood Indices and Human Error Probability Estimates) presents thehuman reliability estimates for each task produced via MAUD during the SLIM-MAUD test application and the direct estimation procedures.

REFERENCES

Comer, M. K., Kozinsky, E. J., Eckel, E. J., and Miller, D. P. (1983), "Human Reliability Data Bank for Nuclear Power Plant Operations, Volume 2: A Data Bank Concept and System Description," NUREG/CR-2744, Sandia National Laboratories, Alburquerque, NM.

Comer, M. K., Seaver, D. A., Stillwell, W. G., and Gaddy, C. D. (1984), "Generating Human Reliability Estimates Using Expert Judgment, Volume 1: Main Report," NUREG/CR-3688, Sandia National Laboratories, Alburquerque, NM.

Embrey, D. E. and Hall, R. E. (1981), "Quantification of Human Performance Shaping Factors," Proceedings of the ANS/ENS Optical Meeting on Probabilistic Risk Assessment, Portchester, NY, LaGrange Park, IL, American Nuclear Society.

Embrey, D. E (1983), "The Use of Performance Shaping Factors and Quantified Expert Judgment in the Evaluation of Human Reliability: An Initial Appraisal," NUREG/CR-2986, Brookhaven National Laboratory, Upton, NY.

Embrey, D. E., Humphreys, P., Rosa, E. A., Kirwan, B., and Rea, K. (1984a) "SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgment, Volume 1: Overview of SLIM-MAUD," NUREG/CR-3518, Brookhaven National Laboratory, Upton, NY.

Embrey, D. E., Humphreys, P., Rosa, E. A., Kirwan, B., and Rea, K. (1984b), "SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgment, Volume 2: Detailed Analysis of the Technical Issues," NUREG/CR-3518, Brookhaven National Laboratory, Upton, NY.

Gustafson, D., Peterson, P., Koetsky, E., von Koningsveld, R., Macco, A., Casper, S., and Rossmeissl, J. (1983), "A Decision Analytic System for Regulating the Quality of Care in Nursing Homes: System Design and Evaluation," in Humphreys, P. C., Svenson, O., and Vari, A. (eds.) <u>Analysis and Aiding Decision Process</u>, Amsterdam: North Holland.

Kruskal, J. B. and Wish, M. (1978), "Multidimensional Scaling," Sage University, paper series on Quantitative Applications in the Social Sciences, 07-011, Beverly Hills, CA: Sage Publications.

Kruskal, J. B., Young, F. W., and Seery, J. B. (1973), "How to use KYST, a Very Flexible Program to do Multidimensional Scaling and Unfolding," Murry Hill, NJ: Bell Laboratories, unpublished.

Nemiroff, P. M. and King, D. C. (1975), "Group Decision Making Performance as Influenced by Consensus and Self-Orientation," <u>Human Relations</u>, 28, 1.

Phillips, L. D. Humphreys, P., and Embrey, D. E. (1985), "Appendix D: A Socio-Technical Approach to Assessing Human Reliability (STAHR)" in Selby, D., Pressurized Thermal Shock Evaluation of the Calvert Cliffs Unit 1 Nuclear Power Plant, Research Report on DOE Contract 105840R21400, Oak Ridge National Laboratory, Oak Ridge, TN.

Phillips, L. D. and Embrey, D. E. (1985), "Appendix E: Quantification of Operator Actions by STAHR," in Selby, D., Pressurized Thermal Shock Evaluation of the Calvert Cliffs Unit 1 Nuclear Power Plant, Research Report on DOE Contract 105840R21400, Oak Ridge National Laboratory, Oak Ridge, TN.

Seaver, D. A. and Stillwell, W. G. (1983), "Procedures for Using Expert Judgment to Estimate Human Error Probabilities in Nuclear Power Plant Operations," NUREG/CR-2743, Sandia National Laboratories, Alburquerque, NM.

Swain A. D. and Guttmann, H. E. (1983), "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications," NUREG/CR-1278, Sandia National Laboratories, Alburquerque, NM.

APPENDIX A

USER MANUAL FOR IMPLEMENTING SLIM-MAUD

A.1.0 INTRODUCTION

A.1.1 Purpose

This manual provides instructions for using the SLIM-MAUD methodology in order to obtain human error probabilities (HEPs) for nuclear power plant tasks. Figure A.1 shows an outline of the procedures in this manual.

A.1.2 What is SLIM-MAUD?

SLIM-MAUD is a set of procedures designed to assist expert judges in estimating HEPs for nuclear power plant tasks for applications in safety studies and risk assessments. SLIM, which stands for the Success Likelihood Index Methodology, (SLIM) elicits and organizes experts' judgments about the factors which influence human performance of nuclear power plant tasks. MAUD, which stands for Multi-Attribute Utility Decomposition (MAUD), is a computer program used to implement SLIM.

The key underlying assumption of SLIM is that the success or failure of any human task depends upon the combined effects of a small set of Performance Shaping Factors (PSFs) which influence a particular task. PSFs include situational characteristics, instructions, task and equipment characteristics, organismic factors, and psychological and physiological stressors. The MAUD program helps experts to identify and rate PSFs which influence a set of tasks, determines the relative importance of the PSFs, and calculates the Success Likelihood Index (SLI) for each task. These SLIs are then converted to HEPs.

A.1.3 Resource Requirements

A.1.3.1 Human Resources

SLIM-MAUD was developed to be used by a group of four experts with a wide range of experience with nuclear power plant tasks and operations. The group may consist of more than four members; however, the minimum size of the group is three members. Group members should be selected who have relevant experience with the tasks to be assessed. Each group should contain at least one individual experienced with design and operations of the specific type of nuclear power plant being assessed, at least one individual with probabilistic risk assessment (PRA) experience, and at least one individual with human factors experience.

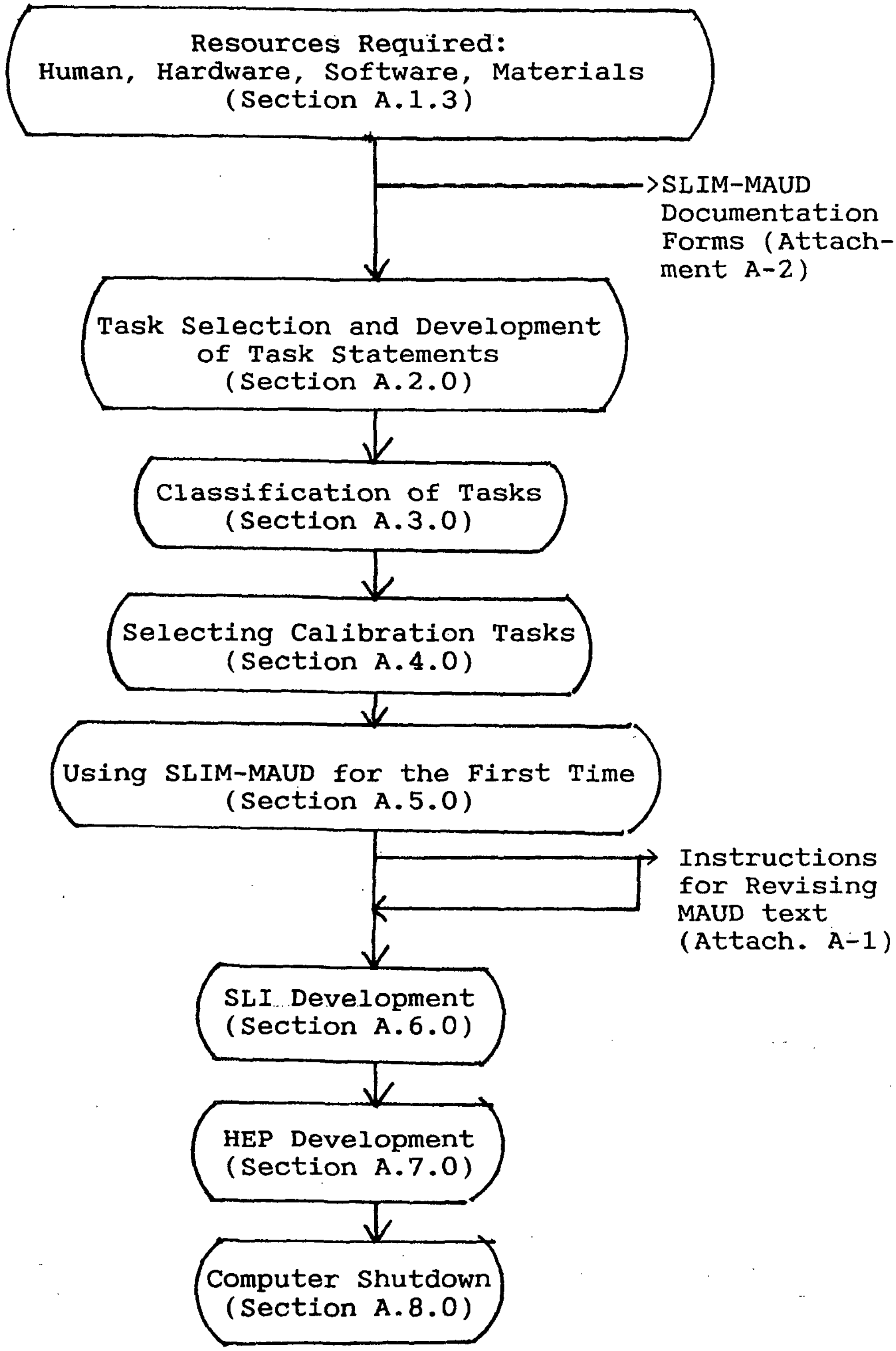


Figure A.1 Outline of procedures for using SLIM-MAUD.

A.1.3.2 Hardware Resources

The hardware needed to implement SLIM-MAUD include: (1) an IBM/PC personal computer with at least 64K Random Access Memory (RAM), monitor, and two 360K disk drives, or an IBM/PC compatible personal computer which runs under CP/M or IBM/PC DOS operating systems and (2) a printer.

A.1.3.3 Other Materials

1. The MAUD Software Diskette

The MAUD software diskette is proprietary to the Decision Analysis Unit of the London School of Economics. To obtain this diskette, an End User License Agreement for the MAUD software program must be purchased from the Decision Analysis Unit. To order MAUD, call or write to the Decision Analysis Unit at:

Decision Analysis Unit of the
London School of Economics
Houghton Street
London WC2A 2AE, England
Telephone: 44-01-405-7686 Telex: 24655 BLPES G

The MAUD program developed by the Decision Analysis Unit is a general-purpose program designed to be applicable in a wide range of decision-making and assessment situations. This general-purpose program can be "reconfigured" or customized for specific applications such as SLIM-MAUD. The reconfiguration involves making changes in the vocabulary used by the program. Instructions for reconfiguring MAUD for a SLIM-MAUD application are contained in Attachment A-1. The U.S. Nuclear Regulatory Commission (NRC) will supply a MAUD diskette reconfigured for a SLIM-MAUD application to individuals who have purchased an End User License for MAUD from the Decision Analysis Unit. Contact:

Division of Automated Services
Office of Research Management
U.S. Nuclear Regulatory Commission
Washington, DC 20555

2. DOS system diskette.

3. Two blank 5-1/4 inch floppy diskettes. A backup copy of the SLIM-MAUD programs will be stored on one diskette. The other diskette will be used to store the data from a SLIM-MAUD session.

4. Index cards (5 x 8 inch) for the development of task statements and for sorting tasks into groups during the task classification process.

5. Photocopies of the forms in Attachment A-3. We recommend that these forms be placed in a three-ring binder so that forms can be easily added, removed and replaced.

6. Pens, pencils, scratch paper.

A.1.4 Organization of this Manual

This manual contains seven sections and two attachments:

Section A.2, Task Selection and Development of Task Statements, describes procedures for task selection and the development of task statements.

Section A.3, Classification of Tasks, describes how to classify the set of tasks into groups for use with the MAUD program.

Section A.4, Selecting Calibration Tasks, describes how to select pairs of reference tasks to be used in developing HEPs.

Section A.5, Starting the SLIM-MAUD program, describes how to start the computer, load the SLIM-MAUD programs, and enter the names of the tasks to be assessed.

Section A.6, SLI Development, describes how to use the MAUD5 program to generate SLIs by identifying, rating, and weighting PSFs and how to interpret the printed summary report.

Section A.7, HEP Development, describes how to convert SLIs into HEPs and obtain Uncertainty Bounds for the HEPs.

Section A.8, Ending the SLIM-MAUD session, describes how to end the SLIM-MAUD session and turn off the computer.

Attachment A-1 contains instructions for revising the text in the MAUD program.

Attachment A-2 contains forms for the documentation of the task selection, task classification, and PSF identification processes which should be copied and collated to make a Documentation Booklet.

A.2.0 TASK SELECTION AND DEVELOPMENT OF TASK STATEMENTS

A.2.1 Purpose

This section provides instructions for selecting tasks and developing detailed task statements that describe the nuclear power plant tasks to be assessed using SLIM-MAUD. These statements must be detailed and concrete so that all members of the group share the same interpretation of the task.

A.2.2 Materials

1. A list of nuclear power plant tasks to be assessed.
2. Task Analysis Forms from Attachment A-2 (one per task to be assessed).
3. SLIM-MAUD Group Members Form from Attachment A-2.
4. 5 x 8 inch index cards.
5. Paper, pencils or pens.

A.2.3 Procedures

1. At the start of each SLIM-MAUD session, a SLIM-MAUD GROUP Members Form should be completed. If group membership changes during the SLIM-MAUD session, a second form should be completed. All completed forms should be retained.

2. Complete a Task Analysis Form for each task. The purpose of this form is to elicit a full verbal description of the task from group members so that the actions and personnel required, and the PSFs influencing task performance are clear to everyone. Directions for filling out this form are provided below.

   a. DATE
   Fill in the current date.

   b. TASK NUMBER
   Assign a number to each task for record keeping purposes. Number tasks consecutively as each form is completed.

   c. TASK LABEL
   Develop a unique label to refer to each task during the SLIM-MAUD session. Try to keep this label short--one to two words if possible--so that the same label can be used to refer to the task during the SLIM-MAUD session.

   d. JOB CATEGORIZATION
   Specify the individual or members of the team who perform the task, (e.g., shift supervisor, senior reactor operator, shift technical advisor).

e. TASK GOAL
State the goal of the task being assessed in terms of desired system state or purpose of performing the task (e.g., Reopen Main Steam Line Non-Return Valves to Establish Condensor Heat Sink).

f. TASK ELEMENTS
Many tasks are comprised of task elements (i.e., a set of actions that must be performed by personnel to achieve the task goal. List all elements that comprise the task, by specifying the action which must be performed, the individual who performs the action, and the specific equipment, instrument, display, control, etc. which is involved (e.g., Auxiliary Operator observes pump discharge pressure meter on panel C05A to verify appropriate discharge pressure). Table A.1 lists actions which may be required during nuclear power plant tasks. However, this list is not exhaustive. Feel free to identify actions not on the list.

Table A.1 Human Action Verbs Applicable to Nuclear Power Plant Tasks

Equipment/Machine Operation			
Open/Close	Position	Maintain	Calibrate
Fill/Drain	Use	Repair	Test
Start/Stop	Adjust	Select	Check
Cognitive Processes			
Monitor	Detect	Calculate	Categorize
Compute	Encode	Extrapolate	Identify
Interpolate	Interpret	Itemize	Read
Recall	Learn	Remember	Tabulate
Translate	Analyze	Choose	Compare
Decide	Diagnose	Estimate	Plan
Predict	Schedule	Design	Recognize
Supervision			
Advise	Verify	Manage	Inspect
Direct	Instruct	Supervise	
Communication/Social Processes			
Write	Tell	Discuss	Transmit
Read	Ask	Confer	Communicate

g. TASK LOCATION
Identify the location(s) where the task is performed.

h. TIME AVAILABLE
Indicate the amount of time available to perform the task and the criteria used to determine this time frame.

i. TASK CHARACTERISTICS/EVENT DESCRIPTION
Identify the task and/or task elements as procedural or cognitive and whether the task is performed during nominal, routine operating conditions or during off-nominal, post-accident conditions by checking the appropriate option. Briefly describe the event for off-nominal, post-accident conditions.

j. TASK INITIATORS
Identify the stimuli, prompts, or signals that would cause the task to be initiated (e.g., annunciators, oral instructions, procedures).

k. JOB PERFORMANCE AIDS
Specify any job performance aids, procedures (oral, written, checklists), instructions, etc. which are used to perform the task.

A.2.4 Documentation

Copy the Task Number (Item B on the Task Analysis Form) and the Task Label (Item C on the Task Analysis Form) onto an index card. Write a description of the task to be performed on the index card, including the information from Items D-K on the Task Analysis Form. The index cards with the Task Descriptions will be used in the Task Classification Procedures. Make sure a SLIM-MAUD Group Members Form has been completed.

A.2.5 Completion

The process of task selection is completed when task statements for all the tasks for which HEPs are needed have been written on index cards. The group can then go on to the next step, Task Classification (Section A.3.0).

An example of a completed Task Analysis Form is presented in Table A.2.

Table A.2 Example of a Completed Task Analysis Form

SLIM-MAUD DOCUMENTATION FORM

TASK ANALYSIS FORM

A. DATE: 3/6/86 B. TASK NO.: 1

C. TASK LABEL:

SPECIFY THE FOLLOWING INFORMATION FOR EACH TASK. (See Section A.2.0 for additional instructions.)

D. JOB CATEGORIZATION(S): Reactor Operator

E. TASK GOAL: Actuate automatic depressurization system to reduce pressure during IOSP transient before using low-pressure core cooling systems.

F. TASK ELEMENTS:

G. TASK LOCATION: Control Room

H. TIME AVAILABLE: 10 minutes

I. TASK CHARACTERISTICS/EVENT DESCRIPTION:

✓ Procedural ____ Cognitive ____ Both
____ Nominal, routine operating conditions
✓ Off-nominal, post-accident operating conditions
Event Description: IOSP transient with HPCI and RCIC inoperable.

J. TASK INITIATORS:

K. JOB PERFORMANCE AIDS:

A.3.0 CLASSIFICATION OF TASKS

A.3.1 Purpose

This section provides instructions for classifying the task statements into groups of tasks which share common PSFs. PSFs are characteristics of individuals, the task, work environment, and situation which affect human performance in a complex man-machine system such as a nuclear power plant (see Table A.3).

A.3.2 Materials

1. Index cards containing task statements.
2. Completed SLIM-MAUD Group Members Form.
3. Task Classification Form from Attachment A-2.
4. Scrap paper for making PSF labels.

A.3.3 Procedures for Task Classification

A.3.3.1 Sorting Task Statements into Groups

The index cards containing the task statements must be sorted into groups. This classification is based upon the experts judgments' about whether tasks share certain PSFs. The experts should use the following procedures for classification:

1. Select one index card containing a task statement.

2. Read the task statement.

3. Discuss the task statement and identify an important PSF for the performance of that particular task. Table A.3 presents a list of PSFs which may influence the performance of nuclear power plant tasks. When a PSF has been identified, write the PSF on a small slip of paper. Put the index card containing the task statement on the table above the slip of paper containing the PSF label. This creates the first group of tasks.

4. Next, examine each of the task statements on the remaining index cards and decide whether the card belongs in the first pile. That is, decide whether the PSF which defines the first group of tasks is relevant to the performance of that task. For example, training may be an important factor influencing the performance of two tasks. These tasks should be placed in the same pile. However, a third task may require few steps and there may be explicit written procedures available to direct the operator. For this task, the group may decide that training is not an important factor. The index card containing this task statement should be placed in a separate pile and identified by a PSF label, e.g., "Procedures."

Table A.3 Performance Shaping Factors

- Job and Task Instructions: Includes written procedures, written and oral instructions and communications, cautions and warnings, plant policies, shop practices, work methods.

- Task and Equipment Characteristics: Includes man-machine interface factors (design of equipment, job aids, tools, fixtures) control-display relationships, instrumentation, team structure and communication patterns, availability of feedback on human performance, task criticality, frequency and repetitiveness of task, perceptual requirements of task, workload, information load, complexity of task, motor requirements of task.

- Situational characteristics: Includes characteristics of the work environment (temperature, humidity, air quality, radiation, lighting, noise, vibration, cleanliness), architectural features, staffing/manning parameters, organizational structure (responsibility, authority, communication channels), actions by other personnel, work schedules (hours of work, work breaks, shift rotation), rewards, recognition, incentives, benefits, promotions.

- Psychological Stressors: Includes stress-related factors such as suddenness of onset, task speed, task load, perceived risk, threats of failure and loss of job, monotonous, degrading, or meaningless work, long, uneventful vigilance periods, conflicting motives of job performance (e.g., accuracy vs. speed) distractions (noise, glare, movement, display flicker, display color).

- Physiological Stressors: Includes duration of stress, fatigue, pain or discomfort, hunger or thirst, temperature extremes, movement constriction, disruption of circadian rhythm.

- Organismic Factors: Includes previous training, experience level, state of current practice or skill, personality and intelligence variables, motivation and attitudes toward work, emotional state, mental or bodily tension or stress, knowledge of required performance standards, physical condition, group identification.

The goal of the task classification is to identify common general factors which influence performance. A PSF may be rated differently for two tasks, but still be relevant to the performance of each. For example, the experts may identify "amount of time available" as a factor influencing the performance of two tasks. The success criteria for one task may be 30 seconds; for the second task it may be 2 hours. Although the amount of time available differs, this factor may be an important determinant of success. Therefore, these tasks should be placed in the same pile labeled "Time Available."

At the end of the task classification process there should be a relatively few number of groups of task. Each group must contain at least four (4) and no more than eight (8) task statements. Therefore, it may be necessary to combine groups or reclassify certain tasks. Since a number of PSFs influence most tasks, combination or reclassification should not be too difficult.

A.3.4 Documentation

Photocopy a Task Classification Form from Attachment A-2 for each group that has been formed. Complete items A, B, and C at this time. Group numbers are to be assigned consecutively.

If group membership has changed, complete a new SLIM-MAUD Group Members Form.

A.3.5 Completion

When each task statement has been classified into groups and a Task Classification Form filled out for each group, the experts may go on to Section 4.0.

A.4.0 SELECTING CALIBRATION TASKS

A.4.1 Purpose

Two tasks must be included in each group of tasks which are similar to the tasks already in the group and for which HEPs are already available. This pair of tasks in each group will be referred to as "calibration reference tasks" because they will serve as reference tasks for calibrating or transforming the SLIs into HEPs for other tasks in the group. The purpose of this section is to explain how to select calibration tasks for each group of task statements.

A.4.2 Materials

1. Calibration Task Analysis Form from Attachment A-2.
2. Completed Task Classification Forms.
3. Completed SLIM-MAUD Group Members Form.

A.4.3 Procedure

Select and define two calibration reference tasks for each group of tasks. Criteria for selecting each pair of calibration tasks are the following:

1. Each pair of calibration reference tasks should be influenced by the PSFs which influence the other tasks in the group.

2. An HEP estimate from an independent source must already be available for each calibration reference task. Sources of HEP data for calibration reference tasks include: NUREG/CR-1278 (Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications), PRAs for other nuclear power plants, and direct estimation of HEPs by experts.

3. The tasks selected for each pair of calibration tasks should represent the broadest range of HEPs possible for tasks sharing similar PSFs. That is, one task should have a relatively high probability of failure and one task should have a relatively low probability of failure.

A.4.4 Documentation

Photocopy and complete a Calibration Task Analysis Form from Attachment A-2 for each calibration task selected. Complete items D, E, and F on the Task Classification Form.

Complete a SLIM-MAUD Group Members Form if group membership has changed.

A.5.0 USING SLIM-MAUD FOR THE FIRST TIME

A.5.1 Purpose

This section provides instructions for (1) copying the SLIM-MAUD files from the master diskette to a "backup" diskette and (2) checking that the SLIM-MAUD program is properly configured for a SLIM-MAUD session.

Please Note: In the instructions below <CR> means to press the Enter key on the keyboard.

A.5.2 Materials

1. IBM/PC or IBM/PC compatible personal computer with at least 64K RAM, monitor, and two 360K disk drives.
2. Printer.
3. DOS system diskette.
4. SLIM-MAUD master diskette.
5. One 5-1/4 inch floppy diskette.
6. Disk Operating System (DOS) manual.

A.5.3 Procedures

A.5.3.1 Prepare a SLIM-MAUD Backup Diskette

Follow the instructions below to copy the SLIM-MAUD programs from the original "master" diskette to a "backup" diskette. It is recommended that the SLIM-MAUD backup diskette be used during SLIM-MAUD sessions and that the original SLIM-MAUD master diskette be stored in a safe place. If the backup diskette becomes damaged, the original diskette can be used to prepare another backup diskette.

1. To format the diskette in Drive B:

   Place your DOS system diskette in Drive A and the blank 5-1/4 inch diskette in Drive B. Turn on the computer. When DOS is ready, the screen will display the DOS ready prompt: A>

   Type: FORMAT B:/S <CR>

   The screen will display the instructions:

   'Insert new diskette for drive B:
   and strike any key when ready.'

   Strike any key on the keyboard.

   When formatting is complete, the screen will display the question:

'Format another (Y/N)?'

Type: N

2. Make sure DOS is ready and A> is displayed.

   Type: COPY BASIC.* B:

3. When copying is finished and A> appears, remove the DOS system diskette from Drive A.

   Place a write-protect tab over the notch on the original SLIM-MAUD master diskette and insert this diskette in Drive A.

4. Make sure DOS is ready and A> is displayed (see note below).

   Type: COPY M*.* B: <CR>

   When the A> is displayed, type: COPY *.BAS B: <CR>

   When the A> is displayed, type: COPY SESSION B: <CR>

   When the A> is displayed, type: COPY INIT B: <CR>

   Note: Instead of using four COPY commands to copy the SLIM-MAUD files to Drive B, you may choose to use the COPY *.* B: command to copy the entire contents of the original diskette onto the diskette in Drive B. However, make sure that the version of DOS on the original SLIM-MAUD diskette is the same as the version on your DOS system diskette, or you may have some problems when running SLIM-MAUD. (See the DOS manual for the VER command.)

5. Make sure DOS is ready and A> is displayed.

   Type: DIR B: <CR>

   The directory of the files on the diskette in Drive B will be displayed. The following files should be on the diskette in Drive B:

   COMMAND.COM
   BASIC.COM
   MAUDOVR6.EXE
   MDSENS6.EXE
   MDMAIN6.EXE
   MAUD.EXE
   MDPREF6.EXE
   MDMOVE.EXE
   MDCONFIG.EXE
   MDPARM.DAT
   MDTEXT.DAT
   HEP.BAS
   SLIPROB.BAS
   SESSION
   INIT

If the diskette in Drive B contains the HEP.BAS file, skip to Section A.5.3.2. If the HEP.BAS file is not on the diskette in Drive B, go on to Step 7.

7. Remove the original SLIM-MAUD diskette from Drive A and place it in a safe place. Switch the SLIM-MAUD backup diskette from Drive B to Drive A. Make sure DOS is ready and A> is displayed. Type:

   BASIC <CR>

   When the screen displays the following BASIC prompt: Ok

   Type each of the following lines and press <CR> at the end of each line.

```
10 'HEP.BAS PROGRAM
20 'CONVERTS SLI VALUES TO HEPS
30 FOR I=1 TO 24 : PRINT : NEXT
40 CLS
50 PRINT "*****************************************************************"
60 PRINT "*                    HEP CALCULATION PROGRAM                    *"
70 PRINT "*                                                               *"
80 PRINT "*    This program calculates Human Error Probabilities (HEPs) *"
90 PRINT "*    for nuclear power plant tasks which have been assessed   *"
100 PRINT "*    using SLIM-MAUD.                                          *"
110 PRINT "*                                                              *"
120 PRINT "*    The inputs to this program are:                           *"
130 PRINT "*    1)  The HEPs associated with each pair of calibration     *"
140 PRINT "*        reference tasks.                                      *"
150 PRINT "*    2)  The Success Likelihood Indices (SLIs) produced        *"
160 PRINT "*        by the SLIM-MAUD program for the pair of calibration *"
170 PRINT "*        tasks.                                                *"
180 PRINT "*    3)  The SLIs produced by the SLIM-MAUD program for each  *"
190 PRINT "*        of the other tasks assessed by SLIM-MAUD.             *"
200 PRINT "*                                                              *"
210 PRINT "*    The outputs of this program are HEPs for each of the      *"
220 PRINT "*    tasks assessed by SLIM-MAUD.                              *"
230 PRINT "*                                                              *"
240 PRINT "*    This program must be rebooted for each group of tasks     *"
250 PRINT "*    to be assessed.                                           *"
260 PRINT "****************************************************************"
270 LINE INPUT; "PRESS THE RETURN KEY TO CONTINUE";XS
271 CLS
272 PRINT "****************************************************************"
273 PRINT "*                                                              *"
274 PRINT "*     This program must be restarted for each group of tasks   *"
275 PRINT "*     assessed using SLIM-MAUD.  The appropriate values for    *"
276 PRINT "*     calibration tasks for each group must be entered into    *"
277 PRINT "*     the program each time the program is restarted.          *"
278 PRINT "*                                                              *"
279 PRINT "*     When HEPs for all the tasks within a group have been     *"
280 PRINT "*     calculated, stop the program by typing N to the question *"
281 PRINT "*    'Is there another HEP to be calculated?'    Then type     *"
282 PRINT "*     RUN to restart the program.                              *"
283 PRINT "****************************************************************"
284 LINE INPUT; "PRESS THE RETURN KEY TO CONTINUE";ZS
285 CLS
290 PRINT "Type in the name of the first calibration reference task"
300 LINE INPUT ">",R1S: PRINT
310 PRINT "What is its HEP";: INPUT P1
320 PRINT "What is its SLI value";: INPUT SLI1:PRINT
330 PRINT "Type in the name of the second calibration reference task"
340 LINE INPUT ">",R2$: PRINT
350 PRINT "What is its HEP";: INPUT P2
360 PRINT "What is its SLI value";: INPUT SLI2: PRINT
370 LINE INPUT "Are these values O.K.?",QS
380 IF QS="" THEN 370
390 QS=LEFTS(QS,1)
400 IF QS="N" THEN 30
410 IF QS="n" THEN 30
420 '
430 'COMPUTER PARAMETER VALUES
440 'COMPUTR PARAMETER-VALUES
450 PRINT
460 PRINT:PRINT
470 CLS
480 P1=LOG(P1): P2=LOG(P2)
490 A=(P1-P2)/(SLI1-SLI2)
500 B=P1-A*SLI1
510 '
520 PRINT "****************************************************************"
530 PRINT "*                                                              *"
540 PRINT "*     Next, you will enter the names and the SLIs of the       *"
550 PRINT "*     nuclear power plant tasks for which HEPs are needed.     *"
560 PRINT "*                                                              *"
570 PRINT "****************************************************************"
580 LINE INPUT; "PRESS THE RETURN KEY TO CONTINUE        ";YS
590 PRINT:PRINT "Type in the name of the task requiring an HEP."
600 LINE INPUT ">",RS:PRINT
610 PRINT "What is its SLI value";: INPUT SLI
620 PRINT:PRINT
630 P=A*SLI+B: P=EXP(P)
640 PRINT "HEP is ";P
650 PRINT "Please write down the HEP calculated for this task."
660 PRINT:PRINT
670 LINE INPUT; "Is there another HEP to be calculated (Y/N)  ?";QS
680 IF QS="" THEN 670
690 QS=LEFTS(QS, 1)
700 IF QS="Y" THEN 590
710 IF QS="y" THEN 590
720 END
```

To save this program file onto the SLIM-MAUD backup diskette in Drive A

Type: SAVE "HEP.BAS" <CR>

To return to DOS and the A>

Type: SYSTEM <CR>

A.5.3.2 Checking That the SLIM-MAUD Program is Properly Configured

Before using SLIM-MAUD for the first time, check to see that it is properly configured.

1. Make sure that DOS is ready and A> is displayed. If the SLIM-MAUD backup diskette is not in Drive A, switch it from Drive B to Drive A. Type:

   MDCONFIG <CR>

```
What would you like to do:

        (1) Change the drive address for MAUD5 sessions
            [Currently set to DRIVE B]

        (2) Revise the text used within MAUD5
        (3) No changes for now

Please type your option number:
```

Read Item 1 on the screen. If the drive address for MAUD5 sessions is "Currently set to DRIVE B," return to DOS by typing:

3 <CR>

If the drive address for MAUD5 sessions is not "Currently set to DRIVE B," type:

1 <CR>

The following questions will be presented on the screen. Answer them as shown below.

```
Current drive address for MAUD5 sessions: B

Do you want to specify a new address? Y

Enter new drive address [A-Z]: B

Is this OK? Y
```

When the menu appears again, type:

3 <CR>

to save the change and return to DOS.

2. When the A> is displayed, type:

   MDCONFIG <CR>

   When the menu shown above appears on the screen, type:

   2 <CR>

   The following text will appear on the screen:

```
You can

        (1) Revise the current text
        (2) Restore the default text
        (3) Revise the default text
        (4) Revise nothing for now

Which would you like to do?
Please type your option number:  1
```

Type: 1 <CR>

If the following text does not appear, go to Attachment A-1 for instructions on how to reconfigure the text used in the SLIM-MAUD programs.

```
MAUD5 starts with a 15-line introductory frame.
At present, it appears like this;

>       SLIM-MAUD is set up on this computer to help you
>to assess the Likelihood of Success of tasks performed
>by personnel in nuclear power plants.
>
>       SLIM-MAUD works in the following way:
>  o    You will be asked to name at least 4 tasks.
>  o    You must identify at least 2 Performance
>       Shaping Factors (PSFs) that you believe affect
>       the likelihood of success or failure of the tasks.
>  o    You must rate each PSF for each task.
>  o    SLIM-MAUD will guide you in judging the relative
>       importance of PSFs and will then calculate weights
>       for PSFs.
>  o    Finally, SLIM-MAUD will calculate Success Likelihood
>       Indices (SLIs) and print out a Summary Report.

Do you wish to change this?
```

If the text shown above <u>does</u> appear on the screen, press the Ctrl key and type "C" at the same time to terminate program execution.

A.5.4 Completion

After following the above procedures, you will have a SLIM-MAUD backup diskette to be used during SLIM-MAUD sessions. Instructions for running the SLIM-MAUD programs for an application are found in the following section (Section 6.0, SLI Development).

A.6.0 SLI DEVELOPMENT

A.6.1 Purpose

This section provides instructions for running the SLIM-MAUD programs to obtain SLIs for tasks within a single group. Directions for starting the program, entering the task labels, identifying PSFs, and interpreting the summary report are contained in this section. These procedures must be repeated for each group of tasks.

A.6.2 Materials

1. SLIM-MAUD Backup Diskette.
2. DOS system diskette.
3. A 5-1/4 inch floppy diskette.
4. IBM/PC computer, printer, monitor.
5. Completed Task Analysis Forms, Calibration Task Analysis Forms, Task Classification Forms, and SLIM-MAUD Group Members Form.
5. Photocopies of Summary of SLIM-MAUD Session Form from Attachment A-2.

A.6.3 Procedures

A.6.3.1 Starting a SLIM-MAUD Session

1. Insert the DOS system diskette in Drive A and turn on computer. Make sure DOS is ready and A> is displayed.

2. Insert a blank 5-1/4 inch floppy diskette in Drive B.

   Type: FORMAT B: <CR>

   The screen will display the following instructions:

   'Insert new diskette for Drive B:
   and strike any key when ready'

   Strike any key on the keyboard.

   When formatting is complete, the screen will display the question:

   'Format another (Y/N)?'

   Type: N

3. Remove the DOS system diskette from Drive A: and insert the SLIM-MAUD backup diskette in Drive A. Make sure DOS is ready and A> is displayed.

   Type: MAUD <CR>

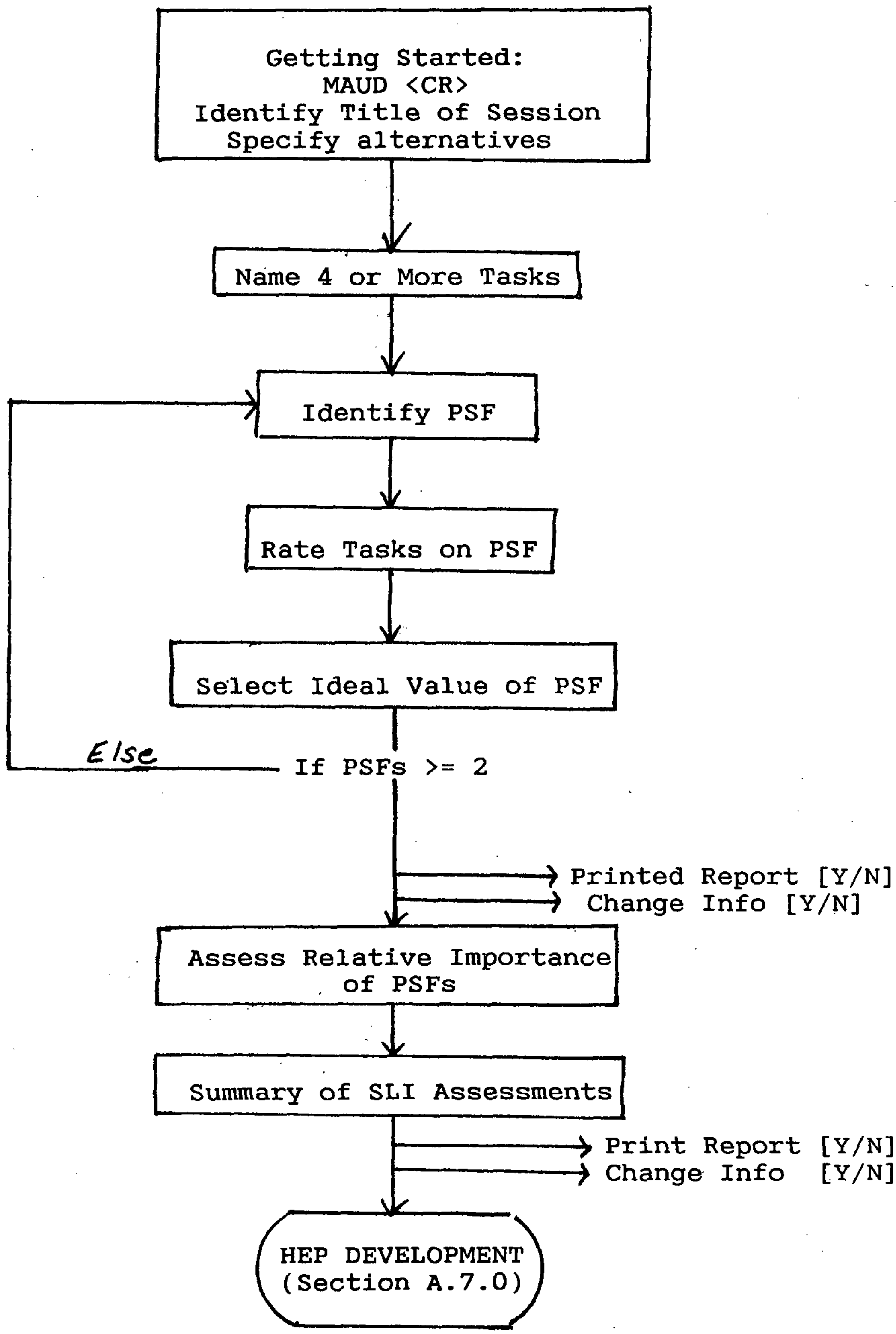


Figure A.2 Flow chart of SLIM-MAUD session.

This command starts the interactive MAUD program. Information will be presented on the computer monitor. Read each display and type in your responses. The following pages of this manual illustrate the operation of the MAUD program. Instructions regarding user responses appear below each display.

A.6.3.2 How the SLIM-MAUD Program Works

Frame 1

- -

```
     SLIM-MAUD is set up on this computer to help you
to assess the Likelihood of Success of tasks performed
by personnel in nuclear power plants.

     SLIM-MAUD works in the following way:
  o  You will be asked to name at least 4 tasks.
  o  You must identify at least 2 Performance
     Shaping Factors (PSFs) that you believe affect
     the likelihood of success or failure of the tasks.
  o  You must rate each PSF for each task.
  o  SLIM-MAUD will guide you in judging the relative
     importance of PSFs and will then calculate weights
     for PSFs.
  o  Finally, SLIM-MAUD will calculate Success Likelihood
     Indices (SLIs) and print out a Summary Report.

Have you used the MAUD system before? n

[ Please type YES or NO, and then press the RETURN key.  If you prefer,
  you can type  Y for 'YES' and  N for 'NO' ]
```

- -

Frame 2

- -

```
MAUD will work with you, using the words you give it.
It will ask you questions which are relevant
in assessing tasks' likelihood of success

It will allow you to make changes in your descriptions of the alternatives
under consideration and your assessments as you wish.

When MAUD asks you a question, for example:

WHAT IS THE TITLE FOR THIS SESSION?

                    SLIM-MAUD SESSION 1

You should type in your answer from the keyboard.
Enter your answer now, and then press the RETURN key
```

- -

Frame 3

- -

The title for this session will be SLIM-MAUD SESSION 1

If you make a typing mistake, there are two ways of making corrections:

- If you notice the mistake before you press the RETURN key, you can correct it by pressing the backspace button, which rubs out the characters one at a time. You can then type the correct characters, followed by pressing RETURN.
- If you notice the mistake after you have pressed the RETURN key, you will have to wait until the next opportunity MAUD offers you to make changes. However, these opportunities occur quite frequently.

- -

* Press the SPACE bar to continue *

Frame 4

- -

Please type in a word describing the topic you want to make a decision about by completing the sentence:

The alternatives I am thinking about could all be described as TASKS

[Type one word such as JOBS or CARS or HOUSES and press the RETURN key]

- -

Instructions

For SLIM-MAUD Applications, the alternatives are usually described as "Tasks" (Frame 4) and a single alternative is described as a "Task" (Frame 5).

Frame 5

- -

All the alternatives can be described as TASKS
so, the word to describe each alternative is TASK

[for example, if all the alternatives are JOBS, each alternative can be described as a JOB]

- -

Frame 6

- -

All the alternatives can be described as TASKS
so, the word to describe each alternative is TASK

[for example, if all the alternatives are JOBS, each alternative can be described as a JOB]

Are you reasonably happy with the words you typed? Y

- -

Frame 7

- -

Please think about the TASKS which you wish
to consider in this session.

You will need to name not less than 4 and not more than 10
TASKS . Each name you type in should be quite short - type
just one or two words.
The name you give to each TASK can be whatever you like,
so long as you are clear to what it refers.

You may wish to name not only TASKS which are
under immediate consideration, but others which may be considered later
available, or those which are useful for reference purposes.

- -

* Press the SPACE bar to continue *

Frame 8

- -

Please type in the name of the first TASK
you would like to consider :

Its name is TASK A

Now the next TASK you want to consider :

Its name is TASK B

Now the next TASK you want to consider :

Its name is TASK C

Now the next TASK you want to consider :

Its name is TASK D

- -

Instructions

Type in the names of the first four tasks in the group of tasks being assessed, Press <CR> after each task name. Refer to the Task Classification Forms when entering task names. All tasks must be members of the same group. Keep the name of each task brief; if the name is too long, MAUD will ask you to type in a shorter name.

Frame 9

- -

Is there another TASK you want to consider ? Y

Its name is TASK E

- -

Instructions

After the first four tasks are entered, MAUD will ask if there are any more tasks in the group. If yes, type Y. Press <CR> and type in the next task name. This will continue until you type N (see Frame 10) or until 10 tasks have been entered.

Frame 10

- -

Is there another TASK you want to consider ? Y

- -

Frame 11

- -

The TASKS currently under consideration are :

[1] TASK A
[2] TASK B
[3] TASK C
[4] TASK D
[5] TASK E

- -

Do you want to change anything? [Enter Y/N]

Frame 12

- -

* Press the SPACE bar to continue *

You are now going to be asked about differences between TASKS . Try to think about a particular Performance Shaping Factor where the differences between the TASKS are important to you in assessing tasks' likelihood of success.

For instance, some people consider certain TASKS to be simple while other TASKS are complex and some TASKS are in between.

This is just one example and may not be relevant to you. There are no right or wrong answers; it is important to try to include those TASKS which you think are important in assessing tasks' likelihood of success within the particular situation you are considering.

- -

Frame 13

```
Can you think of one Performance Shaping Factor
in which one of these TASKS

       [1] TASK A
       [2] TASK B
       [3] TASK C

differs from the other two on a particular Performance Shaping Factor that
matters to you in assessing tasks' likelihood of success? [Enter Y/N]  Y

What is the number next to the TASK  that is different? 2
```

Instructions

Refer back to the PSFs identified on the Task Classification Form for this group of tasks. Discuss the three tasks presented in the display in order to identify a single PSF on which to rate the tasks. For example, one task may differ from the other tasks in the quality of the supervision available in the task situation. That is, one task may be well supervised while the other tasks are less well supervised or poorly supervised.

When the group identifies a single PSF which distinguishes the tasks from one another, type Y <CR>.

If the group is unable to identify a single PSF which distinguishes the tasks from one another, type N <CR>. MAUD will present another group of three tasks for the group to consider.

MAUD will continue presenting groups of three tasks to be considered until the group answers Y.

Next, MAUD asks for the number of the task which differs from the other two. Enter the number of the task (from the display) which is different. In this example, we typed "2" to indicate that Task B differed from Tasks A and C on the basis of a PSF.

Frame 14

- -

```
You have said that TASK B  is different from
TASK A  and TASK C .
In not more than three words each time, please describe
the way in which they differ.

First describe TASK B

TASK B  is :  HIGH WORKLOAD
```

- -

Instructions

Think of the PSF you have identified as a scale with two endpoints. A different amount or level of the PSF may be characteristic of each task in the group. Define one endpoint of the PSF scale by typing in how the PSF is characteristic of the task you identified as different.

Keep the description of the PSF brief (not more than three words) and press <CR>.

Frame 15

- -

```
You have said that TASK B  is different from
TASK A  and TASK C .
In not more than three words each time, please describe
the way in which they differ.

First describe TASK B

TASK B  is :  HIGH WORKLOAD

On the other hand,
TASK A  and
TASK C  are :  LOW WORKLOAD
```

- -

Frame 16

```
You have said that TASK B  is different from
TASK A  and TASK C .
In not more than three words each time, please describe
the way in which they differ.

First describe TASK B

TASK B  is :  HIGH WORKLOAD

On the other hand,
TASK A  and
TASK C  are :  LOW WORKLOAD

Are you reasonably happy with this description? [Enter Y/N] Y
```

Instructions

Now define the other endpoint of the PSF scale by typing in how the PSF is characteristic of other tasks in the group.

Again, keep the description brief (not more than three words) and press <CR>.

In this example, we identified the PSF scale as WORKLOAD. One endpoint of this scale is HIGH WORKLOAD. The other endpoint of this scale is LOW WORKLOAD.

Please Note: the PSF scale should be one-dimensional. That is, it should represent different quantities of a single dimension. HIGH STRESS versus LOW STRESS are acceptable endpoints for a PSF scale. On the other hand, VERY COMPLEX versus LOW STRESS are not acceptable endpoints for a PSF scale because they represent two dimensions.

Frame 17

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

It should be possible to give each TASK
a rating from 1 to 9 according to its position on the scale

          HIGH WORKLOAD              to             LOW WORKLOAD
               |------------------------------------------|
                1    2    3    4    5    6    7    8    9

Can you rate the TASKS  on this scale? [Enter Y/N] Y

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

Instructions

This display represents the PSF scale defined above. There are nine positions on the scale. The endpoints are assigned values of one (1) and (9), respectively. You will be asked to assess the position of each task on this scale. If this scale appears adequate for rating the tasks, type Y <CR>. If not, type N and you will be able to start the process of PSF identification again (see Frame 13).

Frame 18

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

It should be possible to give each TASK
a rating from 1 to 9 according to its position on the scale

          HIGH WORKLOAD              to             LOW WORKLOAD
               |------------------------------------------|
                1    2    3    4    5    6    7    8    9

Move the cursor along the scale using the keys marked < and > until you find
the point which is your assessment of  TASK B
                         and then type a *
-----------------------------------------------------------------------------
Your rating of  1 :TASK A  is                                        3.8

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

Instructions

Use the arrow keys to move the cursor along the scale until you reach the point on the scale which describes the level of the PSF for the task being assessed. Then press the * key. In the example above, the goal was to quantify the level of workload for Task A.

Frame 19

```
It should be possible to give each TASK
a rating from 1 to 9 according to its position on the scale

            HIGH WORKLOAD              to              LOW WORKLOAD
                    |-------------------------------------------|
                    1    2    3    4    5    6    7    8    9

                    Are these ratings O.K.? [Enter Y/N] Y
-----------------------------------------------------------------------
Your rating of  1 :TASK A  is                                     3.8
Your rating of  2 :TASK B  is                                     7.0
Your rating of  3 :TASK C  is                                     8.2
Your rating of  4 :TASK D  is                                     2.0
Your rating of  5 :TASK E  is                                     5.0
```

Instructions

When you press the * key, the numerical value of the assessment for the task appears on the display. In this example, Task A was rated as 3.8 on the WORKLOAD scale.

Next, the level of WORKLOAD for Task B was determined by moving the cursor and pressing the * key. Continue this process for all the tasks in the group.

When all the tasks in the group have been rated on the scale, type Y <CR> or <CR> to indicate whether the ratings are acceptable. In this example, we typed Y <CR>. If you press N <CR>, you will be given the opportunity to make the following changes: (1) cancel this PSF and all ratings on this scale, (2) change a rating on scale, or (3) change nothing for now.

Frame 20

```
Thinking only about the Performance Shaping Factor below, what position
on the scale would be IDEAL for a TASK
in assessing tasks' likelihood of success?

            HIGH WORKLOAD              to              LOW WORKLOAD
                    |-------------------------------------------|
                    1    2    3    4    5    6    7    8    9

Move the cursor along the scale using the keys marked < and > until you find
the location of your best point on this scale, and then type a *
```

Instructions

Move the cursor along the scale to indicate the ideal or optimum value of the PSF for tasks in this group. Type *.

Frame 21

- -

```
Thinking only about the Performance Shaping Factor below, what position
on the scale would be IDEAL for a TASK
in assessing tasks' likelihood of success?

          HIGH WORKLOAD                to              LOW WORKLOAD
                |                                               |
                |-----------------------------------------------|
                1    2    3    4    5    6    7    8    9

Your rating of the ideal point on this scale is 5.0
--------------------------------------------------------------------


Are you happy with all your current ratings on this scale? [Enter Y/N] Y
```

- -

Instructions

In this example, 5.0 was identified as the ideal or optimum value of WORKLOAD.

If the group is satisfied with their ratings of tasks on the PSF scale, and their rating of the ideal PSF value for this scale, type Y <CR>.

If the group wishes to change their ratings of tasks on the PSF scale, type N <CR>. MAUD will then give you the opportunity to: (1) cancel the PSF and all ratings on it, (2) change your ratings on this PSF, (3) change the position of the IDEAL task, or (4) change nothing now.

When all tasks have been assessed on the first PSF scale, and the IDEAL value of the PSF for the group of tasks has been identified, MAUD will present another group of 3 tasks for consideration. Follow the procedures above to identify a second PSF scale and rate the tasks on this scale.

Note: At least two PSF scales must be identified for each group of tasks in order for MAUD to go on to the next phase of SLIM-MAUD (i.e., calculating relative weights for the PSF scales). If these scales are not identified, MAUD will continue presenting triads of tasks as in Frame 22.

Frame 22

- -

```
Can you think of one Performance Shaping Factor
in which one of these TASKS

     [1] TASK C
     [2] TASK D
     [3] TASK E

differs from the other two on a particular Performance Shaping Factor that
matters to you in assessing tasks' likelihood of success? [Enter Y/N]  Y
                      * Please type Y or N *
```

- -

Frames 23-32 illustrate how two more PSF scales were developed: (1) WELL SUPERVISED to POORLY SUPERVISED and (2) GOOD PROCEDURES to POOR PROCEDURES.

Frame 23

- -

```
You have said that TASK C  is different from
TASK D  and TASK E .
In not more than three words each time, please describe
the way in which they differ.

First describe TASK C

TASK C  is :  WELL SUPERVISED

On the other hand,
TASK D  and
TASK E  are :  POORLY SUPERVISED

Are you reasonably happy with this description? [Enter Y/N] Y
```

- -

Frame 24

- -

```
It should be possible to give each TASK
a rating from 1 to 9 according to its position on the scale

          WELL SUPERVISED            to          POORLY SUPERVISED
                 |----------------------------------------|
                  1    2    3    4    5    6    7    8    9

                    Are these ratings O.K.? [Enter Y/N] y
```

Your rating of	1 :TASK A is	1.0
Your rating of	2 :TASK B is	3.0
Your rating of	3 :TASK C is	4.2
Your rating of	4 :TASK D is	7.0
Your rating of	5 :TASK E is	8.2

- -

Frame 25

- -

```
Thinking only about the Performance Shaping Factor below, what position
on the scale would be IDEAL for a TASK
in assessing tasks' likelihood of success?

          WELL SUPERVISED            to          POORLY SUPERVISED
                 |----------------------------------------|
                  1    2    3    4    5    6    7    8    9

Your rating of the ideal point on this scale is 1.0


Are you happy with all your current ratings on this scale? [Enter Y/N] y
```

- -

Frame 26

Would you like to have a printed report of the information
you have put in so far? [Enter Y/N] N

Frame 27

Do you want to change any of the information you
have put in so far? [Enter Y/N] N

Frame 28

Do you want to specify any other relevant Performance Shaping Factors
in which the TASKS differ from each other
in a way that matters in assessing tasks' likelihood of success? [Enter Y/N] Y

In not more than three words each time, please describe
how some of them differ from the others

Some TASKS are : GOOD PROCEDURES

Whereas other TASKS are : POOR PROCEDURES

Frame 29

It should be possible to give each TASK
a rating from 1 to 9 according to its position on the scale

```
GOOD PROCEDURES              to              POOR PROCEDURES
       |-------------------------------------------|
        1     2     3     4     5     6     7     8     9
```

Are these ratings O.K.? [Enter Y/N] Y

Your rating of	1 :TASK A is	7.0
Your rating of	2 :TASK B is	5.8
Your rating of	3 :TASK C is	3.4
Your rating of	4 :TASK D is	2.4
Your rating of	5 :TASK E is	1.0

Frame 30

- -

Thinking only about the Performance Shaping Factors below, what position on the scale would be IDEAL for a TASK in assessing tasks' likelihood of success?

```
GOOD PROCEDURES                to              POOR PROCEDURES
                |-----------------------------------------|
                 1    2    3    4    5    6    7    8    9
```

Your rating of the ideal point on this scale is 1.0

- -

Are you happy with all your current ratings on this scale? [Enter Y/N] Y

Frame 31

- -

You can

(1) ASSESS the relative importance of each Performance Shaping Factors and then see your overall assessments for the TASKS .
(2) CHANGE some of the information you have put in.
(3) ADD another Performance Shaping Factors.
(4) PRINT a report of the session so far.
(5) SAVE your data for future use.
(6) STOP.

Which would you like to do?

Please type a number: 1

- -

Instructions

When all PSFs have been identified, Frames 26 and 27 will appear again. When these questions are answered, the menu in Frame 31 will appear. Type '1' <CR> to go on to the next phase of SLIM-MAUD.

Frame 32

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
In order to find out which are the most important factors for you
in assessing tasks' likelihood of success, you will now have to choose
between hypothetical TASKS  which MAUD
will describe to you, using some of your own words.

Would you like some instructions on how to make your choice? Y
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

Instructions

The purpose of the next phase of MAUD is to find out how the experts weight the relative importance of the PSFs for determining the likelihood of success for each task. If you have not used SLIM-MAUD before and/or are unfamiliar with this process, type Y <CR> to get additional information.

In this example, we typed Y.

Frame 33

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

You are going to be asked to choose between 2 hypothetical
TASKS  which differ on just two of your scales.

Let's look at the first TASK

TASK  A scores as follows :

        POOR PROCEDURES              to          GOOD PROCEDURES
           |                                            |
           |----------*-------------------------------|

        LOW WORKLOAD                 to          HIGH WORKLOAD
           |                                            |
           |-------------------*----------------------|

- On the first scale, the  *  indicates that TASK  A
  scores the same as the WORST TASK  you rated on this scale.

- On the second scale, the  *  indicates that TASK  A
  scores the same as your IDEAL TASK  on this scale.
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

* Press the SPACE bar to continue *

Frame 34

- -

Let's look at the second TASK

TASK B scores as follows :

```
          POOR PROCEDURES              to              GOOD PROCEDURES
                 |--------------------------------------*|

            LOW WORKLOAD               to               HIGH WORKLOAD
                 |----*-----------------------------------|
```

- On the first scale, the * indicates that TASK B scores the same as your IDEAL TASK on this scale.
- On the second scale, the * indicates that TASK B scores the same as the WORST TASK you rated on this scale.

- -

* Press the SPACE bar to continue *

Instructions

The experts will be asked to make more of these comparisons. When MAUD presents such tasks, the experts must consider how the tasks have been defined, and which task is more likely to succeed. That should be the basis for choosing or "preferring" a task.

Frame 35

- -

* Press the SPACE bar to continue *

These two hypothetical TASKS will be presented to you together, and you will be asked which one you assess as most likely to succeed

If you choose TASK A, its rating on the second scale will be moved a little way down the scale to make a new, slightly less attractive hypothetical TASK A.

If, on the other hand, you choose TASK B, then its rating on the first scale will be moved a little way down the scale to make a new, slightly less attractive hypothetical TASK B.

You will then be asked again to choose between the two TASKS until MAUD finds the point at which your preference between TASKS A and B changes over.

- -

* Press the SPACE bar to continue *

Instructions

Frames 35 - 40 illustrate how MAUD works to elicit the relative importance weights of the PSFs, based on the experts' decisions.

Frame 35 illustrates how a hypothetical task scores on two PSF scales. This task is characterized by relatively POOR PROCEDURES and IDEAL WORKLOAD. Compare hypothetical Task B in Frame 35. Task B is characterized by IDEAL PROCEDURES and LOW WORKLOAD (which is not ideal for successful task performance).

Frame 36

- -

* Press the SPACE bar to continue *

```
This process will be repeated 2 times, using hypothetical
TASKS  described on various pairs of Performance Shaping Factorss
after which MAUD will know enough about which factors are important to you
in choosing between TASKS  to work out your
assessments for the
REAL TASKS
which you have been considering up till now.
```

- -

Instructions

In Frames 37-38, tasks were compared on the basis of their ratings on Procedures and Workload. In Frames 39-40, tasks were compared on the basis of their ratings on Procedures and Supervision. The SLIM-MAUD users typed A or B each time they were asked to make a change. Each time a choice was made, the asterisk denoting a task's rating moved on one scale to make one of the tasks more or less desirable. Then the SLIM-MAUD users were asked to again choose between Task A or B until their perference for one of the tasks changed (see Frames 37 and 40).

Frame 37

- -

```
Imagine you had to choose between TASK  A
which scores as follows :

           POOR PROCEDURES             to             GOOD PROCEDURES
                   |                                         |
                   |----------*------------------------------|

            LOW WORKLOAD               to              HIGH WORKLOAD
                   |                                         |
                   |--------------------*--------------------|

and TASK  B which scores as follows :

           POOR PROCEDURES             to             GOOD PROCEDURES
                   |                                         |
                   |-----------------------*-----------------|

            LOW WORKLOAD               to              HIGH WORKLOAD
                   |                                         |
                   |----*------------------------------------|

OK. Now which would you choose, A or B :
```

- -

Frame 38

- -

```
Imagine you had to choose between TASK  A
which scores as follows :

           POOR PROCEDURES             to             GOOD PROCEDURES
                   |                                         |
                   |----------*------------------------------|

            LOW WORKLOAD               to              HIGH WORKLOAD
                   |                                         |
                   |--------------------*--------------------|

and TASK  B which scores as follows :

           POOR PROCEDURES             to             GOOD PROCEDURES
                   |                                         |
                   |-----------------------*-----------------|

            LOW WORKLOAD               to              HIGH WORKLOAD
                   |                                         |
                   |----*------------------------------------|

OK. Now which would you choose, A or B :A
Are you reasonably sure [Y/N]? Y
```

- -

Frame 39

Imagine you had to choose between TASK A
which scores as follows :

```
          POOR PROCEDURES              to              GOOD PROCEDURES
                    |----------*-----------------------------|

         POORLY SUPERVISED             to              WELL SUPERVISED
                    |---------------------------------------*|
```

and TASK B which scores as follows :

```
          POOR PROCEDURES              to              GOOD PROCEDURES
                    |----------------------------*-----------|

         POORLY SUPERVISED             to              WELL SUPERVISED
                    |----*-----------------------------------|
```

OK. Now which would you choose, A or B :

Frame 40

Imagine you had to choose between TASK A
which scores as follows :

```
          POOR PROCEDURES              to              GOOD PROCEDURES
                    |----------*-----------------------------|

         POORLY SUPERVISED             to              WELL SUPERVISED
                    |---------------------------------------*|
```

and TASK B which scores as follows :

```
          POOR PROCEDURES              to              GOOD PROCEDURES
                    |----------------------------*-----------|

         POORLY SUPERVISED             to              WELL SUPERVISED
                    |----*-----------------------------------|
```

OK. Now which would you choose, A or B :A
Are you reasonably sure [Y/N]? Y

Instructions

When the weighting process is finished, MAUD will display the relative likelihood of success for each of the tasks assessed during this session (Frame 41). The Success Likelihood Indices (SLIs) will be based on the ratings of each task on the PSF scales and the relative importance of each PSF scale. SLIs can range from 1.00 to 0.00 with 1.00 indicating "most likely to succeed" and 0.00 indicating "least likely to succeed."

Experts can obtain a detailed Summary Report of the session. The Summary Report for this session is shown in Table 6.2. Section A.6.3. contains information about interpreting the Summary Report.

Frame 41

- -

```
Your order of assessments for the TASKS
from best to worst is :
                                        Success Likelihood Indices
TASK E                                          (0.70)  < BEST
TASK C                                          (0.45)
TASK A                                          (0.45)
TASK D                                          (0.42)
TASK B                                          (0.40)  < WORST


------------------------------------------------------------------------
      Would you like to have a printed report of the information
      you have put in so far? [Enter Y/N] Y
```

- -

Instructions

Examine the relative SLIs for each task in the group and the ratings of the tasks on each PSF scale shown on the report. If the experts would like to change their assessments, type Y <CR>.

In this example, we typed N.

Frame 42

```
Your order of assessments for the TASKS
from best to worst is :
                                          Success Likelihood Indices
TASK E                                           (0.70)  < BEST
TASK C                                           (0.45)
TASK A                                           (0.45)
TASK D                                           (0.42)
TASK B                                           (0.40)  < WORST
```

Do you want to change any of the information you have put in so far? [Enter Y/N] N

Instructions

If the group would like to rate the tasks on another PSF, type Y <CR>. In Frame 43, we typed N.

Frame 43

```
Do you want to specify any other relevant Performance Shaping Factor
in which the TASKS  differ from each other
in a way that matters in assessing tasks' likelihood of success? [Enter Y/N] N
```

Frame 44

```
You can

      (1) ASSESS the relative importance of each
          Performance Shaping Factors and then see your overall
          assessments for the TASKS .
      (2) CHANGE some of the information you have put in.
      (3) ADD another Performance Shaping Factors.
      (4) PRINT a report of the session so far.
      (5) SAVE your data for future use.
      (6) STOP.
      (7) EXAMINE how your overall assessments for the
          TASKS  depends on the relative importance
          of any particular Performance Shaping Factors.

Which would you like to do?
Please type a number: 5
```

Instructions

If the results of the SLIM-MAUD session are satisfactory, type "5" to save the data from this session. Make sure the previously-formatted SLIM-MAUD data diskette is in Drive B.

Frame 45

- -

```
The information you have entered so far will be stored on the disk in drive B.

You must specify a file name which can be up to 8 characters long, starting
with an alphabetic character.  You can then use any combination of
alphabets or numbers to complete the name.  The maximum limit is
8 characters.


Please type the name of the file in which you want to keep the material
from this session: SLIS1

Is this name O.K.? Y
```

- -

Instructions

Type in a name for the data file for this session. In this example we named the data file "SLIS1." This file will be stored on the SLIM-MAUD data diskette in drive B as SLIS1.MD. All SLIM-MAUD data files are automatically assigned the ".MD" file extension, therefore do not include this in the name of the data file.

After the data file is written to Drive B, the MAUD menu appears again (see Frame 44). Type "6" to end this SLIM-MAUD session.

A.6.3.3 Restarting MAUD to Assess Another Group of Tasks

MAUD must be restarted to assess another group of tasks. To restart MAUD, make sure DOS is ready and A> appears. Type:

MAUD <CR>

Follow the instructions in Section A.6.3.2.

A.6.3.4 Using a SLIM-MAUD Data File from a Previous Session

A SLIM-MAUD session may be started using a data file from a previous SLIM-MAUD session. The experts may choose to do this if they want to add or delete tasks, change ratings of tasks on PSFs, or change ratings of the relative importance of PSFs. Any of these changes will result in a new set of SLIs for the tasks in question.

To start a SLIM-MAUD session using data from a previous session:

1. Insert the SLIM-MAUD backup diskette in Drive A. Make sure DOS is ready and A> appears.

2. Insert the SLIM-MAUD data diskette containing the file you want to use in Drive B.

3. Type MAUD <CR>.

4. Frame 46 will appear on the display. Type Y to indicate that you have used MAUD before.

Frame 46

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
**********************************<MAUD6>**********************************
---------------------------------------------------------------------------
      SLIM-MAUD is set up on this computer to help you
to assess the Likelihood of Success of tasks performed
by personnel in nuclear power plants.

      SLIM-MAUD works in the following way:
  o   You will be asked to name at least 4 tasks.
  o   You must identify at least 2 Performance
      Shaping Factors (PSFs) that you believe affect
      the likelihood of success or failure of the tasks.
  o   You must rate each PSF for each task.
  o   SLIM-MAUD will guide you in judging the relative
      importance of PSFs and will then calculate weights
      for PSFs.
  o   Finally, SLIM-MAUD will calculate Success Likelihood
      Indices (SLIs) and print out a Summary Report.

Have you used the MAUD system before? Y

[ Please type YES or NO, and then press the RETURN key.  If you prefer,
  you can type  Y for 'YES' and  N for 'NO' ]
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

Next, Frame 47 will appear on the display. Type in the name of the data file in Drive B that you want to use.

N.B. Do not specify the drive location when typing in the name of the data file or MAUD will respond with an error message. That is, B:SLIS1 is incorrect. Instead, use "SLIS1." It is not necessary to specify the file extension. MAUD automatically assigns all SLIM-MAUD data files in the .MD extension. Therefore, if you type in the name of the data file without the .MD extension, MAUD automatically assumes the extension is .MD.

Frame 47

```
Are you starting this MAUD session from scratch? N
(rather than starting with data on file from a previous session)

Please type in the name of the file you want to work with.

         Its name is:  SLIS1
```

Next, Frame 48 will appear on the display.

Frame 48

```
You can

      (1) ASSESS the relative importance of each
          Performance Shaping Factor and then see your overall
          assessments for the TASKS .
      (2) CHANGE some of the information you have put in.
      (3) ADD another Performance Shaping Factor.
      (4) PRINT a report of the session so far.
      (5) SAVE your data for future use.
      (6) STOP.
      (7) EXAMINE how your overall assessments for the
          TASKS  depends on the relative importance
          of any particular Performance Shaping Factor.

Which would you like to do?
Please type a number:
```

If you choose Option 1, the experts will have the opportunity to re-weigh the relative importance of the PSF scales used in the previous assessment (see Frames 33 to 40). SLIM-MAUD will calculate new SLIs for the tasks.

If you choose Option 2, Frame 49 will appear on the display. Here, the experts have the opportunity to change the tasks that were assessed or change the existing PSF scales. If tasks or PSFs are changed, the experts will then have to rerate the tasks on each PSF scale and re-weigh the relative importance of the PSFs (see Frames 22 and 40). Again, SLIM-MAUD will calculate new SLIs for the tasks.

Option 2 would be particularly useful for examining how improvements in PSFs might affect the tasks' likelihood of success. The experts might assume that training or procedures were improved for certain tasks. They could then re-rate the tasks on the relevant PSF scales. MAUD would then calculate new SLIs for the tasks. This exercise would assist the experts in evaluating whether proposed improvements would be a cost-effective way of improving human reliability.

Frame 49

- -

```
You can

       (1) Change the set of TASKS  under consideration
           by adding, deleting or changing TASKS
       (2) Change some Performance Shaping Factor on which you have rated the
           TASKS
       (3) Change nothing now

Which would you like to do?
Please type the number :
```

- -

Option 3 (in Frame 48) allows the experts to add another PSF scale. Then the tasks in the group will have to be rated on the new PSF scale and the experts will have to re-weigh the relative importance of the PSFs.

If Option 4 is selected, the experts can obtain a printed report of the SLIM-MAUD session.

Option 5 allows the experts to save the results of the SLIM-MAUD session in the data file on the diskette in Drive B. This was shown in Frames 44 and 45.

Option 7 performs a sensitivity analysis of the ordering of the tasks on the SLI scale based on varying the ratings of each task on each PSF. Frames 50 through 52 illustrate how this sensitivity analysis is performed.

Frame 50

- -

The order of preferences of the TASKS
may vary according to the relative importance (or `weight') assigned to each
Performance Shaping Factor on which they are rated.

The following analysis will allow you to see how the preference ordering of
the TASKS you have assessed
is sensitive to the amount of importance assigned to a
particular Performance Shaping Factor,
while the pattern of the relative importances over every other
Performance Shaping Factor remains unchanged.

The importance weight of the Performance Shaping Factor that
you choose to examine will be varied in 10 equal steps over the whole range:
from 0.0, where it has no importance at all (that is, it plays no part
in determining the preference ordering),
to 1.0, where it completely determines the preference ordering on its own.

At each step, the position of each TASK
in the rank order of preferences will be displayed,
where 1 = best (first in rank) and 3 = worst (last in rank).

- -

* Press the SPACE bar to continue *

Frame 51

- -

The Performance Shaping Factor currently in use are :

1 : HIGH WORKLOAD to LOW WORKLOAD
2 : WELL SUPERVISED to POORLY SUPERVISED
3 : GOOD PROCEDURES to POOR PROCEDURES
What is the index number of the Performance Shaping Factor
whose relative importance you wish to vary? 1

- -

Frame 52

- -

The table below shows the effect on the rank ordering of
assessments for the TASKS of varying the relative importance of
the Performance Shaping Factor scaled from HIGH WORKLOAD
to LOW WORKLOAD :

Name of TASK	Order of preference (1 = best, 5 = worst)										
TASK A	5	4	3	2	2	2	2	2	2	2	2
TASK B	4	5	5	4	3	3	3	3	3	3	3
TASK C	2	2	2	3	4	5	5	5	5	5	5
TASK D	3	3	4	5	5	4	4	4	4	4	4
TASK E	1	1	1	1	1	1	1	1	1	1	1
Possible range of relative importance >>	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0

of this Performance Shaping Factor(0 = no importance, 1.0 = total importance)
(Note that the current ACTUAL relative importance value is 0.23)

- -

Frame 52 shows the results of the sensitivity analysis when the relative importance of the HIGH WORKLOAD - LOW WORKLOAD scale was varied. This frame shows how the rank ordering of the tasks in terms of Success Likelihood would change if the PSF were rated as more or less important by the experts.

A.6.4 Documentation

At the end of each SLIM-MAUD session, a Summary Report is printed. Figure A.1 contains an example of a Summary Report. The circled numbers on the report refer to the comments below.

Use the information in the Summary Report to complete the Summary of SLIM-MAUD Session Form from Attachment A-2. Copy the tasks numbers, labels, and respective SLIs onto this form.

If group membership has changed, complete a new SLIM-MAUD Group Member Form.

A.6.5 Completion

At the end of each SLIM-MAUD session, the experts will have SLIs for each task within the group of tasks just assessed. The experts should have a printed Summary Report of the SLIM-MAUD session, a partially filled out Summary of SLIM-MAUD Session Form, and have saved the data from the current session onto the diskette in Drive B. At this point, another group of tasks can be assessed or the experts can go onto Section A.7.0.

A.6.6 Understanding the SLIM-MAUD Summary Report

Figure A.1 contains a copy of the summary report for the session illustrated in the above example. The sections of the report have been numbered in the left margin of the table. These circled numbers refer to the items discussed below.

1. The first line of the report contains the title of the session given by the group at the start of the SLIM-MAUD session.

2. Section 2 is a table which presents the ratings given to each task on each 9 point PSF rating scale.

   a. The names of the tasks assessed in this session are listed vertically. In this example, the tasks were labeled TASK A through TASK E.

   b. The number of each PSF rating scale is shown to the left of the ratings. HIGH WORKLOAD to LOW WORKLOAD is #1; WELL SUPERVISED to POORLY SUPERVISED is #2; GOOD PROCEDURES to POOR PROCEDURES is #3.

   c. The name of each PSF rating scale is shown on the right side of the table. The name of each PSF scale are the two endpoints of the scale. Thus, the first rating scale is "HIGH WORKLOAD(1) to LOW WORKLOAD(9)." The number in parenthese after each endpoint label indicates that on this 9 point scale, "1" represented the HIGH WORKLOAD end of the scale, "9" represented the LOW WORKLOAD end of the scale.

(1.) SUMMARY OF SESSION SLIM-MAUD SESSION 1 SO FAR:

(2) The ratings of the TASKS on the scales you are currently using are as follows:

(2a) (2b) (2c)

Rating, plus ideal value scale number	TASK A	TASK B	TASK C	TASK D	TASK E	Name of Performance Shaping Factor and relative importance
(1)	3.8	7.0	8.2	2.0	5.0	HIGH WORKLOAD (1) to LOW WORKLOAD (9) * Ideal value= 5.0 (2d) * Relative importance = 0.23 (2e)
(2)	1.0	3.0	4.2	7.0	8.2	WELL SUPERVISED (1) to POORLY SUPERVISED (9) * Ideal value= 1.0 * Relative importance = 0.30
(3)	7.0	5.8 (2f)	3.4	2.4	1.0	GOOD PROCEDURES (1) to POOR PROCEDURES (9) * Ideal value= 1.0 * Relative importance = 0.47

(3) On the basis of the information shown above, the current order of assessments for the TASKS (from best to worst) should be as follows:

TASK E (0.70) <BEST

TASK C (0.45)

TASK A (0.45)

TASK D (0.42)

TASK B (0.40) <WORST

The numerical values, shown in parentheses above, index the strengths of your assessments.
A value of 1.0 would indicate that a TASK is best on every Performance Shaping Factor;
a value of 0.0 would indicate that a TASK is worst on every Performance Shaping Factor.

Figure A.3 Summary of SLIM-MAUD Session 1, 11/25/85 so far.

(4)

Your order of assessments for the TASKS
from best to worst is :

	Success Likelihood Indices	
TASK E	(0.70)	< BEST
TASK C	(0.45)	
TASK A	(0.45)	
TASK D	(0.42)	
TASK B	(0.40)	< WORST

---

(5)

* The following information shows your assessments
for the TASKS under consideration.

1.0 represents the best TASK and 0.0 represents the
worst TASK on each Performance Shaping Factor.

Rating scale number	TASK A	TASK B	TASK C	TASK D	TASK E	Name of Performance Shaping Factor and relative importance
(1)	0.62	0.38	0.00	0.06	1.00	HIGH WORKLOAD to LOW WORKLOAD * Relative importance = 0.23
(2)	1.00	0.72	0.56	0.17	0.00	WELL SUPERVISED to POORLY SUPERVISED * Relative importance = 0.30
(3)	0.00	0.20	0.60	0.77	1.00	GOOD PROCEDURES to POOR PROCEDURES * Relative importance = 0.47

---

(6) Any TASK which has a score on each Performance Shaping Factor
which is equal to, or higher than, the score of
a particular TASK is said to DOMINATE
that TASK . That is, it should always be preferred, regardless
of the relative importance assigned to each Performance Shaping Factor.

In the table printed above,
No TASK dominates any other TASK
and so the final assessments ordering depends on the
relative importance you assign to each Performance Shaping Factor.

END OF SUMMARY

Figure A.3 Summary of SLIM-MAUD Session 1, 11/25/85 so far (continued).

d. The "ideal value" on the PSF scale is shown below the name of the scale on the right side of the table. The "ideal value" refers to the position on the scale which the experts identified as "IDEAL for a task in assessing tasks' likelihood of success." For the HIGH WORKLOAD to LOW WORKLOAD scale, the ideal point on the scale was 5.0.

e. The relative importance of each scale, in relation to other PSF scales used in the assessment process is given. The relative importance weight for each PSF is calculated by MAUD based on the group's preference choices for hypothetical tasks (see Section 6.2 above, Frames 48-52). In this example, the HIGH WORKLOAD to LOW WORKLOAD scale had a relative weight of 0.23. WELL SUPERVISED to POORLY SUPERVISED was weighted 0.30. The GOOD PROCEDURES to POOR PROCEDURES scale had a relative weight of 0.47. Therefore, the PSF GOOD PROCEDURES to POOR PROCEDURES was weighted somewhat more heavily than the other scales when the SLIs were calculated. The experts should examine the relative importance weights assigned to each PSF to ensure that they accurately reflect their opinion.

f. The body of the table contains the actual numerical ratings given to each task by the group. Thus, the order of tasks in terms of amount of workload (from HIGH to LOW) is:

TASK D	2.0
TASK A	3.8
TASK E	5.0
TASK B	7.0
TASK C	8.2

Note that TASK E was rated as occupying the IDEAL position on this scale.

In terms of WELL SUPERVISED to POORLY SUPERVISED, the tasks were rated as follows:

TASK A	1.0
TASK B	3.0
TASK C	4.2
TASK D	7.0
TASK E	8.2

Task A occupied the ideal position on this scale.

In terms of GOOD PROCEDURES to POOR PROCEDURES, the tasks were rated as follows:

TASK E	1.0
TASK D	2.4
TASK C	3.4
TASK B	5.8
TASK A	7.0

TASK E was assigned the ideal rating on this scale.

3. Section 3 of the summary report contains the SLIs for each task. SLIs may range from 1.0 to 0.0. Tasks are arranged in order from BEST to WORST (i.e., from "most likely to succeed" to "least likely to succeed"). SLIs are calculated from the ratings each task received on the PSFs in relation to the ideal position on the scale and the relative importance weight of each PSF.

   The tasks were ordered from most likely to succeed to least likely to succeed as follows:

TASK E	0.70
TASK C	0.45
TASK A	0.45
TASK D	0.42
TASK B	0.40

   Thus, TASK E was assessed as more likely to succeed than any of the other tasks. This is due to the fact that both workload and procedures were rated as ideal for TASK E.

   The experts should examine the relative SLIs calculated for each task in relation to the PSFs ratings they assigned to tasks, in order to confirm that the ratings "make sense" and are in accord with their understanding of the tasks.

4. Section 4 contains the same information on relative SLI orderings contained in Section 3.

5. Section 5 is a table which contains standardized ratings of the tasks on the PSF scales used. That is, rather than showing the actual ratings each task received on the nine point PSF scale, the ratings have been standardized on a scale from 0.00 to 1.00 where 0.00 represents the worst position on the scale and 1.0 represents the ideal position on the scale.

   This format makes it easier to examine the relative position of tasks across PSF scales because the ideal positions of the scales are identical.

6. The final section of the report contains results of a "dominance analysis." This information is not currently used as part of a SLIM-MAUD assessment.

Please Note: SLIs reflect the relative likelihood of success for the tasks assessed within a single session. SLIs for tasks assessed during different sessions can not be compared. Also, SLIs are not simply the inverse of HEPs. That is, an SLI of 1.00 can not be directly converted into an HEP of 0.00. SLIs can be converted to HEPs by calibrating the SLI scale against a scale of HEPs. (See Section A.7.0).

A.7.0 HEP DEVELOPMENT

A.7.1 Purpose

This section provides instructions for using the HEP.BAS program to generate HEPs and uncertainty bounds from the SLIs produced for each subset of tasks assessed.

A.7.2 Materials

1. One "Summary of SLIM-MAUD Session" form for each subset of tasks.
2. SLIM-MAUD backup diskette.
3. IBM/PC computer, monitor, printer.

A.7.3 Procedures

A.7.3.1 Starting the HEP.BAS Program

1. Insert the SLIM-MAUD Backup System Diskette in Drive A. Make sure DOS is ready and A> appears.

2. Type: BASIC HEP <CR>

   This command loads the HEP.BAS program and begins to run the program as shown in Frames 53 through 59 below. Instructions for using the program appear on the screen. Enter the SLIs and HEPs in decimal form.

   If you make an error while using the HEP.BAS program, you may get the message:

   ?Redo from start

   Hold down the Ctrl key and type "C" at the same time. This will terminate the program and the BASIC prompt "Ok" will appear on the screen. Then type: RUN <CR> to restart the program.

   This program only produces output to the CRT and does not save the results of the run on a diskette. Therefore, you must copy down the HEPs for each of the tasks on the "Summary of SLIM-MAUD Session" form as each one appears on the CRT. If you like, you can have everything that appears on the screen printed out by doing the following before running the HEP.BAS program:

   a. Turn the printer on.
   b. Hold down the "Ctrl" key and press the "PrtSc" key.

   Everything that appears on the screen will be echoed to the printer. To turn this command off, hold down the "Ctrl" key again and press "PrtSc" key.

The BASIC prompt "Ok" will appear. To restart the HEP.BAS program for another group of tasks, type "RUN." To return to DOS, type SYSTEM.

Frame 53

```
************************************************************
*                    HEP CALCULATION PROGRAM                *
*                                                           *
*    This program calculates Human Error Probabilities (HEPs) *
*    for nuclear power plant tasks which have been assessed  *
*    using SLIM-MAUD.                                        *
*                                                           *
*    The inputs to this program are:                         *
*    1)  The HEPs associated with each pair of calibration   *
*        reference tasks.                                    *
*    2)  The Success Likelihood Indices (SLIs) produced      *
*        by the SLIM-MAUD program for the pair of calibration *
*        tasks.                                              *
*    3)  The SLIs produced by the SLIM-MAUD program for each *
*        of the other tasks assessed by SLIM-MAUD.           *
*                                                           *
*    The outputs of this program are HEPs for each of the    *
*    tasks assessed by SLIM-MAUD.                            *
*                                                           *
*    This program must be rebooted for each group of tasks   *
*    to be assessed.                                         *
************************************************************
PRESS THE RETURN KEY TO CONTINUE
```

Frame 54

```
************************************************************
*                                                          *
*     This program must be restarted for each group of tasks *
*     assessed using SLIM-MAUD.  The appropriate values for *
*     calibration tasks for each group must be entered into *
*     the program each time the program is restarted.       *
*                                                          *
*     When HEPs for all the tasks within a group have been  *
*     calculated, stop the program by typing N to the question *
*    'Is there another HEP to be calculated?'    Then type  *
*     RUN to restart the program.                           *
************************************************************
PRESS THE RETURN KEY TO CONTINUE
```

Frame 55

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Type in the name of the first calibration reference task
>TASK A

What is its HEP? .07
What is its SLI value? .45

Type in the name of the second calibration reference task
>TASK E

What is its HEP? .001
What is its SLI value? .70

Are these values O.K.?Y
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

Frame 56

```
***************************************************************
*                                                             *
*     Next, you will enter the names and the SLIs of the      *
*     nuclear power plant tasks for which HEPs are needed.    *
*                                                             *
***************************************************************
PRESS THE RETURN KEY TO CONTINUE
```

Frame 57

```
***************************************************************
*                                                             *
*     Next, you will enter the names and the SLIs of the      *
*     nuclear power plant tasks for which HEPs are needed.    *
*                                                             *
***************************************************************
PRESS THE RETURN KEY TO CONTINUE
Type in the name of the task requiring an HEP.
>TASK B

What is its SLI value? .40


HEP is  .163726
Please write down the HEP calculated for this task.


Is there another HEP to be calculated (Y/N)  ?Y
- - - - - - - - - - - - - - - - - - - - - - - - - - -
```

Frame 58

```
***************************************************************
PRESS THE RETURN KEY TO CONTINUE
Type in the name of the task requiring an HEP.
>TASK B

What is its SLI value? .40


HEP is  .163726
Please write down the HEP calculated for this task.


Is there another HEP to be calculated (Y/N)  ?Y
Type in the name of the task requiring an HEP.
>TASK C

What is its SLI value? .45


HEP is  7.000001E-02
Please write down the HEP calculated for this task.


Is there another HEP to be calculated (Y/N)  ?
- - - - - - - - - - - - - - - - - - - - -
```

Frame 59

```
Is there another HEP to be calculated (Y/N)  ?Y
Type in the name of the task requiring an HEP.
>TASK C

What is its SLI value? .45


HEP is  7.000001E-02
Please write down the HEP calculated for this task.


Is there another HEP to be calculated (Y/N)  ?Y
Type in the name of the task requiring an HEP.
>TASK D

What is its SLI value? .42


HEP is  .1165494
Please write down the HEP calculated for this task.


Is there another HEP to be calculated (Y/N)  ?N
```

Run the HEP.BAS program separately for each group of tasks assessed in a SLIM-MAUD session (i.e., for each group of tasks that appears on a SLIM-MAUD summary report). When HEPs have been obtained for the last task in a single group, answer N to the final question in Frame 59:

Is there another HEP to be calculated (Y/N) ?

The BASIC prompt "Ok" will appear on the screen. To restart the HEP.BAS program to obtain HEPs for another group of tasks, type:

RUN <CR>

To exit the HEP.BAS program and return to DOS, type:

SYSTEM <CR>

A.7.3.2 Estimating Upper and Lower Uncertainty Bounds for HEPs

The tasks within a subset share similar characteristics but differ in quality of degree of PSFs. The range of HEPs produced using SLIM-MAUD and HEP.BAS represent a range of probabilities given variations in PSFs. Thus, the tasks with the highest and lowest HEPs in the group represent the upper and lower uncertainty bounds for the HEPs within a group of tasks.

Identify the upper and lower uncertainty bounds on the Summary of SLIM-MAUD Session Form by writing "UB" and "LB" in the column labeled "Uncertainty Bounds" for the tasks with the highest and lowest HEPs.

A.7.4 Documentation and Completion

Make sure a Summary of SLIM-MAUD Session Form has been fully completed for all groups of tasks. Attach all SLIM-MAUD Group Members Forms completed during each session to the Summary of SLIM-MAUD Session Form.

A.8.0 COMPUTER SHUTDOWN

A.8.1 Purpose

The section provides instructions for terminating a SLIM-MAUD session and lists the documentation materials which should be completed.

A.8.2 Ending a SLIM-MAUD Session

Before ending a SLIM-MAUD session, the user may want to make a backup copy of the SLIM-MAUD data diskette. If a backup data diskette is desired, follow Steps 1 through 7 below. If no backup is desired, omit Steps 2 through 5 below.

1. Remove the SLIM-MAUD backup diskette from Drive A.

2. Make sure DOS is ready and the A> appears.

3. Insert a blank formatted diskette in Drive A which will be the SLIM-MAUD backup data diskette. Make sure the original SLIM-MAUD data diskette is still in Drive B.

4. Use the DOS COPY command to copy the data file from Drive B to drive A. This command takes the form:

   COPY B:(filename.MD) A:

   where (filename.MD) refers to the name given to the SLIM-MAUD data file. Do this for each SLIM-MAUD data file on Drive B.

5. Use the DIR A: command to check that all files have been successfully copied onto the diskette in Drive A.

6. Remove the SLIM-MAUD data diskettes from Drives A and B. Label each diskette with the names of the SLIM-MAUD data files.

7. Turn off computer.

A.8.3 Documentation

The following documentation materials should exist at the end of the SLIM-MAUD session.

- SLIM-MAUD data diskette
- SLIM-MAUD backup data diskette
- Completed forms:
  - One Task Analysis Form for each task assessed including calibration tasks.
  - One Task Subset Form for each subset of tasks assessed.
  - One Summary of SLIM-MAUD Session Form and SLIM-MAUD.
  - Group Members Form (attached) for each subset of tasks.

ATTACHMENT A-1

INSTRUCTIONS FOR REVISING MAUD TEXT

1.0 INTRODUCTION

MAUD is usually supplied as a general purpose program for assisting groups or individuals in evaluating alternatives and making decisions. Key words and phrases can be changed in the program so that the questions and text used by MAUD are more suitable for a specific application. Text can be revised by running the program on the SLIM-MAUD diskette called "MDCONFIG."

The text that MAUD is supplied with is called the "default text." When MAUD is used for a SLIM-MAUD application, this text is revised to make it more appropriate for assessing nuclear power plant tasks. The U.S. Nuclear Regulatory Commission (NRC) plans to revise the default text in MAUD to "SLIM-MAUD text" on all SLIM-MAUD diskettes they supply. However, if the SLIM-MAUD diskette you receive does not present "SLIM-MAUD text," use the following procedures to revise the text. Thereafter, the SLIM-MAUD text will be presented every time MAUD is run until the MDCONFIG program is run again.

2.0 PROCEDURES

1. Insert the SLIM-MAUD diskette in Drive A.

2. Make sure DOS is ready and A> appears.

3. Type:

   MDCONFIG <CR>

   The MDCONFIG program runs interactively. The program will present phrases and paragraphs used by the MAUD program and asks if you want to make any changes. Frames 60 through 74 illustrate the changes you should make in order to reconfigure the MAUD program for a SLIM-MAUD application.

Frame 60

```
What would you like to do:

        (1) Change the drive address for MAUD5 sessions
            [Currently set to DRIVE B]

        (2) Revise the text used within MAUD5
        (3) No changes for now


Please type your option number:  2
```

Instructions

Select Option 2 and press the Enter key.

Frame 61

```
You can

        (1) Revise the current text
        (2) Restore the default text
        (3) Revise the default text
        (4) Revise nothing for now

Which would you like to do?
Please type your option number:  1
```

Instructions

Select Option 1 and press the Enter key.

Frame 62

- -

```
MAUD5 starts with a 15-line introductory frame.
At present, it appears like this;

>
>
>
>
>
>
>     This computer is set up to help you think about
>               a decision you wish to make.
>
>
>
>
>
>
>

Do you wish to change this? y
```

- -

Instructions

Type Y and press the Enter key.

Frame 63

- -

```
Please type in your new text ( a maximum of 15 lines)

>        SLIM-MAUD is set up on this computer to help you
> to assess the Likelihood of Success of tasks performed
> by personnel in nuclear power plants.
>
>        SLIM-MAUD works in the following way:
>   o    You will be asked to name at least 4 tasks.
>   o    You must identify at least 2 Performance
>        Shaping Factors (PSFs) that you believe affect
>        the likelihood of success or failure of the tasks.
>   o    You must rate each PSF for each task.
>   o    SLIM-MAUD will guide you in judging the relative
>        importance of PSFs and will then calculate weights
>        for PSFs.
>   o    Finally, SLIM-MAUD will calculate Success Likelihood
>        Indices (SLIs) and print out a Summary Report.

Is this frame O.K?y
```

- -

Instructions

Type in the lines of text shown in Frame 63 exactly as they appear above. Press the Enter key after each line. If you make a typing mistake you may use the backspace key to correct the error if you are on the same line. If you need to correct an error that appears on a previous line, press the Enter key until you have reached the bottom of the display. The question will appear "Do you wish to change this?" Type Y (press the Enter key) and you may begin retyping the introductory frame.

Frame 64

```
Currently, MAUD5 aims to help the user in
>making your decision




Do you wish to change this? y
                              Please type in your new text
>assessing tasks likelihood of success
```

Instructions

Replace "making your decision" with "assessing tasks' likelihood of success."

Frame 65

```
The user is expected to develop his or her
>preferences




Do you wish to change this? y
                              Please type in your new text
>assessments
```

Instructions

Replace "preferences" with "assessments."

Frame 66

- -

```
The alternatives are rated on dimensions, each of which is
currently termed
>aspect


Do you wish to change this? y
                            Please type in your new text
>Performance Shaping Factors (PSFs)
```

- -

Instructions

Replace "aspect" with "Performance Shaping Factors (PSFs)."

Frame 67

- -

```
Currently, the user is asked to define his or her own name
for the set of alternatives under consideration


Do you wish to change this? y
Do you wish the user to define his or her own input? n
```

- -

Instructions

Answer Y to the question "Do you wish to change this?" Then answer N to the question "Do you wish the user to define his or her own input?"

Frame 68

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
      Please supply a new name for the set of alternatives
      In singular form, each alternative is a >task
      In plural form, all alternatives are >tasks

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

Instructions

Enter "task" and press the Enter key. Then enter "tasks" and press the Enter key.

Frame 69

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
        The use is given an example of a aspect
        where some alternatives are characterizes as
        >interesting
        and other alternatives are characterized as
        >boring




        Do you wish to change this? y

        Please supply your new example:
        Some alternatives should be characterized as >simple
        and other alternatives should be characterized as >complex
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

Instructions

Answer Y that you wish to change this frame. Then type "simple" and press the Enter key, and type "complex" and press the Enter key.

Frame 70

```
Numerical values assigned by MAUD5 to alternatives
are currently called
>preference values




Do you wish to change this? y
                            Please type in your new text
>Success Likelihood Indices (SLIs)
```

Instructions

Replace "Preference Values" with "Success Likelihood Indices (SLIs)."

Frame 71

```
The verb describing the basis for choosing an alternative
is to >prefer




Do you wish to change this? y
                            Please type in your new text
>assess as most likely to succeed
```

Instructions

Replace "prefer" with "assess as most likely to succeed."

Frame 72

- -

MAUD5 currently asks for the title for the session with the question
WHAT IS >THE TITLE FOR THIS SESSION

Do you wish to change this? N

- -

Instructions

Do not change this frame.

Frame 73

- -

Are all the changes you have made to the text O.K? Y

- -

Frame 74

- -

What would you like to do:

(1) Change the drive address for MAUD5 sessions
[Currently set to DRIVE B]

(2) Revise the text used within MAUD5
(3) No changes for now

Please type your option number: 3

- -

Instructions

End the MDCONFIG program by selecting Option 3.

ATTACHMENT A-2

SLIM-MAUD DOCUMENTATION FORMS

1. TASK ANALYSIS FORM
2. CALIBRATION TASK ANALYSIS FORM
3. TASK CLASSIFICATION FORM
4. SLIM-MAUD GROUP MEMBERS FORM
5. SUMMARY OF SLIM-MAUD SESSION FORM

ATTACHMENT A-2

1. TASK ANALYSIS FORM

A. DATE:__________ B. TASK NO.:________

C. TASK LABEL:

SPECIFY THE FOLLOWING INFORMATION FOR EACH TASK. (See Section A.2.0 for additional instructions.)

D. JOB CATEGORIZATION(S): ______________________________

____

____

E. TASK GOAL: ______________________________

____

____

F. TASK ELEMENTS: ______________________________

____

____

G. TASK LOCATION: ______________________________

____

____

H. TIME AVAILABLE: ______________________________

____

____

I. TASK CHARACTERISTICS/EVENT DESCRIPTION:

____ Procedural ____ Cognitive ____ Both
____ Nominal, routine operating conditions
____ Off-nominal, post-accident operating conditions
____ Event Description:

J. TASK INITIATORS: ______________________________

____

____

K. JOB PERFORMANCE AIDS: ______________________________

____

____

ATTACHMENT A-2

2. CALIBRATION TASK ANALYSIS FORM

A. DATE:__________ B. TASK NO.:_______

C. TASK LABEL:

SPECIFY THE FOLLOWING INFORMATION FOR EACH TASK. (See Section A.2.0 for additional instructions.)

D. JOB CATEGORIZATION(S): ______________________________

____

____

E. TASK GOAL: ______________________________

____

____

F. TASK ELEMENTS: ______________________________

____

____

G. TASK LOCATION: ______________________________

____

____

H. TIME AVAILABLE: ______________________________

____

____

I. TASK CHARACTERISTICS/EVENT DESCRIPTION:

____ Procedural ____ Cognitive ____ Both
____ Nominal, routine operating conditions
____ Off-nominal, post-accident operating conditions
____ Event Description:

J. TASK INITIATORS: ______________________________

____

____

K. JOB PERFORMANCE AIDS: ______________________________

____

____

ATTACHMENT A-2

3. TASK CLASSIFICATION FORM

A. GROUP #:

B. LIST THE PSFs RELEVENT TO TASKS WITHIN THIS GROUP:

1. ____________________
2. ____________________
3. ____________________
4. ____________________

C. LIST THE TASK NUMBERS AND TASK LABELS (FROM TASK ANLAYSIS FORM) FOR THE TASKS ASSIGNED TO THIS GROUP.

Task #:____ Task Label: ________
Task #:____ Task Label: ________
Task #:____ Task Label: ________
Task #:____ Task Label: ________
Task #:____ Task Label: ________
Task #:____ Task Label: ________
Task #:____ Task Label: ________
Task #:____ Task Label: ________

D. LIST THE TASK NUMBER AND TASK LABEL (FROM TASK ANLAYSIS FORM) FOR CALIBRATION REFERENCE TASK #1 ASSIGNED TO THIS GROUP:

Task No.:____ Task Label: ________
Source of HEP data for this calibration task:

HEP: ______________ Upper Uncertainty Bound: ______________
______________ Lower Uncertainty Bound: ______________

E. LIST THE TASK NUMBER AND TASK LABEL (FROM TASK ANLAYSIS FORM) FOR CALIBRATION REFERENCE TASK #2 ASSIGNED TO THIS GROUP:

Task #: _____ Task Label: ____________
Source of HEP data for this calibration task:

HEP: ______________ Upper Uncertainty Bound: ______________
______________ Lower Uncertainty Bound: ______________

F. TOTAL NUMBER OF TASKS IN GROUP (INCLUDING CALIBRATION TASKS): __________

ATTACHMENT A-2

4. SLIM-MAUD GROUP MEMBERS FORM

Fill out this form at the start of each SLIM-MAUD session and whenever.

GROUP MEMBERSHIP CHANGES:

DATE: ____________________

LOCATION: ____________________

____

____

GROUP MEMBERS

NAME: AFFILIATION/ADDRESS:

1. ____________________

____

Area of Expertise:

2. ____________________

____

Area of Expertise:

3. ____________________

____

Area of Expertise:

4. ____________________

____

Area of Expertise:

5. ____________________

____

Area of Expertise:

ATTACHMENT A-2

5. SUMMARY OF SLIM-MAUD SESSION FORM

Fill out this form at the end of each SLIM-MAUD session.

GROUP NUMBER: ______________

Task No.	LABEL	SLI	HEP	Uncertainty Bounds
Calib. Task 1				
Calib. Task 2				

APPENDIX B

DETAILED METHODS AND RESULTS OF THE TEST APPLICATION OF SLIM-MAUD

This appendix describes the methods and results of the test application of the Success Likelihood Index Methodology - Multi-Attribute Utility Decomposition (SLIM-MAUD) as a technique for estimating human error probabilities (HEPs) in nuclear power plants. The SLIM-MAUD test was evaluated on the basis of three broad criteria: practicality, acceptability, and usefulness. These criteria comprised a comprehensive list of specific issues. These criteria, issues, and the methods and data used to address each are presented in Table B.1. The methods employed, and results of the SLIM-MAUD test application with respect to each of the criteria and issues in Table B.1 will be discussed in this appendix.

Table B.1 SLIM-MAUD Test: Issues and Methods

Issues	Methods/Data
Practicality:	
P1 Cost	Compilation of actual costs incurred.
P2 Subject Matter Experts	Test sessions conducted with groups composed of PRA and human factors experts and individuals with operating experience.
P3 Support Requirements	Specification of equipment and human resources needed.
P4 Transportability	Implementation of test in two locations.
P5 Expandability	Task level compatibility.
P6 Time Requirements	Time expended for each task level.
P7 Interface with Human Reliability Data Bank	Ensured by tasks chosen for evaluation.
P8 Implementability of Procedure	Implementation by minimally trained facilitator.
Acceptability:	
A1 Scientific Community	Submission to professional journals.
A2 Expert Participants	Survey results.
A3 Potential Users	Informal comparative evaluation.
A4 Nuclear Regulatory Commission (NRC)	Not addressed directly (indirect evidence from survey results).
A5 Nuclear Facilities	Not addressed directly (indirect evidence from survey results).
Usefulness:	
U1 Reliability	Consistency of SLI estimates produced.
U2 Face Validity	Survey results.
U3 Convergent Validity	Comparisons with estimates produced by other techniques.

Appendix B is organized in the following manner. First, an overview of the SLIM-MAUD methodology is presented. Next, a detailed description of the evaluation methods used during each stage of the test application is presented. Finally, details of the analyses of the results of the test application, with respect to each of the evaluation criteria, are presented.

B.1.0 OVERVIEW OF SLIM-MAUD

SLIM-MAUD is a methodology which utilizes an interactive computer-based procedure, MAUD, to elicit and organize experts' judgments regarding the factors which affect the likelihood of the successful performance of specific human tasks in nuclear power plants. A detailed description of the theoretical foundations and assumptions of the SLIM-MAUD methodology is contained in Sections 1.11 through 1.13 in Embrey et al. (1984b). Appendix A of this volume provides detailed instructions for implementing this methodology along with illustrations of SLIM-MAUD operations.

The SLIM-MAUD methodology for obtaining experts' estimates of HEPs for nuclear power plant tasks is based on the assumption that the likelihood of successfully performing a task is a function of various characteristics of the individual, situation, and the task. Explicit consideration of these characteristics, known as Performance Shaping Factors (PSFs), is the basic underpinning of SLIM-MAUD. PSFs for nuclear power plant tasks include characteristics of the human such as competence or skill, characteristics of the particular task such as equipment design, procedures, and task complexity, and characteristics of the work environment such as supervision, stress, and time available.

When experts use SLIM-MAUD, the interactive computer program directs them to identify a set of tasks, identify a set of PSFs common to these tasks, rate how good or bad each PSF is for each task, and weight the relative importance of the PSFs. Then the MAUD program uses these judgments of the experts to calculate a Success Likelihood Index (SLI) which represents the relative likelihood of success for each task within the set. The experts can then use a calibration procedure to convert these SLIs to HEPs.

B.2.0 METHODS USED IN THE TEST APPLICATION OF SLIM-MAUD

A plan for the test application of SLIM-MAUD was outlined by Embrey et al. (1984a, Section 8.0, 1984b, Section 4.0). The test plan was designed to enable the utility of the MAUD-based implementation of SLIM to be assessed on the basis of three key criteria: practicality, acceptability, and usefulness. This plan was implemented, with certain revisions, and carried out under NRC contract to the Department of Nuclear Energy, Brookhaven National Laboratory between June and December, 1984.

The principal revisions to the test plan were:

1. Subject matter experts participating in each group evaluated the set of tasks under consideration using both SLIM- MAUD and the direct numerical estimation psychological scaling procedure described by Comer et al. (1984).

2. Four groups of subject matter experts were used, instead of the five in the original test plan.

3. The scope of the test plan analyses of the reliability and validity of SLIM-MAUD was considerably expanded, allowing a much fuller examination of the usefulness of SLIM-MAUD.

B.2.1 Stages in the Test

The test was divided into the following stages:

Stage 1 - Selection of tasks for assessment in the test.

Stage 2 - Classification of tasks into subsets for simultaneous assessment within SLIM-MAUD.

Stage 3 - Selection of the members of the four subject matter expert groups for stage 4.

Stage 4 - Use of SLIM-MAUD by each subject matter expert group for each subset of tasks, followed by direct numerical assessment of all tasks in all subsets by each group member.

Stage 5 - Analysis and interpretation of results from SLIM-MAUD sessions with respect to the issues outlined in Table B-1.

The procedures followed in each of these stages are described below.

B.2.2 Stage 1: Selection of Tasks for Assessment in the Test

For reasons of compatibility, the tasks assessed in the SLIM-MAUD test were identical to those employed in Comer et al.'s (1984) test of psychological scaling methods employing wholistic judgment. Comer et al. (1984)

developed written descriptions of 15 Level A and 20 Level B tasks. Level A tasks corresponded to Level 1 task in the Human Reliability Data Bank (Comer et al., 1983) and Level B tasks corresponded to tasks from Levels 2 and 3 of the Human Reliability Data Bank. Level A tasks involve human-machine interfaces at the systems level: Level B tasks involve human-machine interfaces at the components and instruments/displays/controls levels. Level A tasks are generally more complex than Level B tasks.

The entire set of 15 Level A tasks and 15 of the 20 Level B tasks were selected for use in the test application of SLIM-MAUD. Five Level B tasks were excluded from the test application in order to achieve equal task set size to facilitate comparisons between pairs of tasks and clustering of tasks during the application. Tasks were excluded from the Level B subset if the task description seemed inadequate, incomplete, ambiguous, or if the task appeared to be influenced by a unique set of PSFs. The following Level B tasks were excluded from the test application of SLIM-MAUD.

Task 4. The controls in a control room are all designed so that they are moved to the right if the operator wants to turn on a component. The operator makes an error and turns a rotary control that has three or more positions to the left when he intends to turn the component on.

Task 11. A locally-operated valve has a rising stem and a position indicator. An auxiliary operator, while using written procedures to check a valve lineup, fails to realize that the valve is not in its proper position after a maintenance person has performed a procedure intended to restore it to its proper position after maintenance.

Task 13. An operator incorrectly reads information from a graph that is in a procedure.

Task 16. An operator reads a digital indicator incorrectly.

Task numbers refer to the numbers originally assigned by Comer et al. (1984).

Attachment B-1, Section 1, contains descriptions of Level A and B tasks used in the SLIM-MAUD evaluation.

B.2.3 Stage 2: Classification of Tasks into Subsets for Simultaneous Assessment Within SLIM-MAUD

Tasks assessed using SLIM-MAUD must be grouped into subsets of 4 through 10 tasks which are reasonably homogeneous with regard to the PSFs which are thought to influence performance. Since the final sets of Level A and B tasks covered a wide range of nuclear power plant tasks, it was necessary for the tasks to be classified into subsets in terms of common PSFs.

Classification Procedures

Eight subject matter experts were recruited to participate in a three-stage classification procedure. The subject matter experts were four PRA experts, two human factors experts and two individuals with nuclear power plant operating experience.

B.2.3.1 First Stage - Consensus Over Task Meaning

During the first stage of the task classification procedure, the subject matter experts were each given a booklet containing written instructions and a list of the tasks they would be considering (see Attachment B-1, Section B of this appendix for these instructions). Several days after receiving this booklet, the eight subject matter experts met to discuss the task descriptions, clarify task descriptions, and develop unique abbreviations to refer to each task.

Unique abbreviations were developed for each of the Level A tasks. The eight experts then spent nearly an entire day discussing the meaning of the tasks and the contexts in which such tasks usually take place. They were able to arrive at consensus on each task meaning. In a very few instances the subject matter experts made very minor changes in the wording of the task descriptions.

B.2.3.2 Second Stage - Paired Comparison Rating

The eight subject matter experts were divided into two groups, each consisting of two PRA experts, one human factors expert, and one expert with operating experience. Each group met on separate days to judge the similarity of the important PSFs among tasks. Pairs of task descriptions were presented on viewgraphs. Each group member assigned a number from 1 to 9 reflecting low to high perceived similarities between tasks. Group consensus was achieved through group discussion (see Gustafson et al., 1983; Nemiroff and King, 1975).

Each group was then presented all possible [n(n-1)/2] pairs of tasks. Each pair of tasks was discussed and repeated ratings on the nine-point scale were made until consensus was reached. It took each group approximately one and one-half days to complete the ratings of all tasks.

Figures B.1 through B.4 show the half matrices of task interrelatedness judgments on the nine point scale made by the two groups of subject matter experts for the 15 Level A and 15 Level B tasks. There are no scores in the main diagonal of these half matrices since tasks were not compared with themselves.

Task No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	–														
2	4	–													
3	2	8	–												
4	2	5	6	–											
5	6	3	3	3	–										
6	2	3	5	3	4	–									
7	2	4	5	4	7	2	–								
8	3	2	3	3	2	8	2	–							
9	8	3	2	2	4	2	4	3	–						
10	5	3	6	2	3	3	2	2	3	–					
11	6	6	7	4	3	3	5	2	3	7	–				
12	5	5	4	2	3	4	3	6	3	3	3	–			
13	3	3	3	2	2	6	2	6	3	3	2	4	–		
14	2	2	5	7	3	2	5	2	1	2	5	3	2	–	
15	2	5	6	2	3	7	3	4	3	4	3	6	7	2	–

Task number

Figure B.1 Judged interrelatedness of 15 Level A tasks by Group 1.

Task No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	–														
2	5	–													
3	4	6	–												
4	3	6	7	–											
5	5	3	4	7	–										
6	3	3	4	5	4	–									
7	3	6	5	6	5	7	–								
8	2	3	3	3	3	4	5	–							
9	7	2	2	3	5	2	3	4	–						
10	5	2	2	3	3	4	3	6	7	–					
11	7	2	2	3	6	2	3	3	7	8	–				
12	6	4	3	5	6	5	4	4	4	4	8	–			
13	4	4	3	4	7	4	4	4	5	5	4	6	–		
14	4	5	5	7	4	4	6	2	2	2	1	4	4	–	
15	2	4	4	5	4	6	6	6	2	2	2	4	4	3	–

Task Number

Figure B.2 Judged interrelatedness of 15 Level A tasks by Group 2.

Task No	1	2	3	5	6	7	8	9	10	12	14	15	17	18	19
1	–														
2	5	–													
3	4	5	–												
5	5	3	4	–											
6	3	3	6	3	–										
7	5	4	9	3	7	–									
8	6	6	7	4	4	8	–								
9	9	7	5	5	3	6	6	–							
10	7	6	4	6	3	6	6	8	–						
12	2	2	3	2	2	2	3	3	2	–					
14	2	3	3	2	5	3	3	3	3	2	–				
15	3	4	4	3	3	4	4	3	4	3	3	–			
17	6	5	6	3	6	4	4	4	3	3	3	4	–		
18	7	6	5	4	4	3	5	7	5	4	4	5	7	–	
19	3	3	2	3	2	3	2	3	2	7	2	7	3	4	–

Task numbers

Figure B.3 Judged interrelatedness of 15 Level B tasks by Group 1.

Task No	1	2	3	5	6	7	8	9	10	12	14	15	17	18	19
1	–														
2	7	–													
3	6	7	–												
5	5	6	4	–											
6	3	4	4	6	–										
7	6	7	9	5	5	–									
8	7	6	8	6	4	8	–								
9	9	7	7	6	3	7	8	–							
10	8	7	7	8	5	6	7	8	–						
12	3	4	3	2	2	3	5	4	2	–					
14	2	1	1	1	1	2	1	2	1	2	–				
15	2	2	3	3	3	3	3	3	2	3	7	–			
17	3	4	6	4	4	6	6	6	4	6	2	3	–		
18	4	4	5	5	4	6	6	7	3	7	2	3	8	–	
19	4	4	6	4	5	5	5	6	2	7	2	4	7	7	–

Task numbers

Figure B.4 Judged interrelatedness of 15 Level B tasks by Group 2.

B.2.3.3 Third Stage - Multidimensional Scaling Analysis

Non-metric multidimensional similarities scaling was used to construct two dimensional "task-interrelatedness" maps of the 15 Level A tasks, and the 15 Level B tasks. The procedure used is described in Kruskal (1964a,b), and developed and implemented in a computer program called KYST (Kruskal, Young, and Seery, 1973). Details of the procedures, and the availability of the program are given in Kruskal and Wish (1978) and Schiffman, Reynolds, and Young (1981).

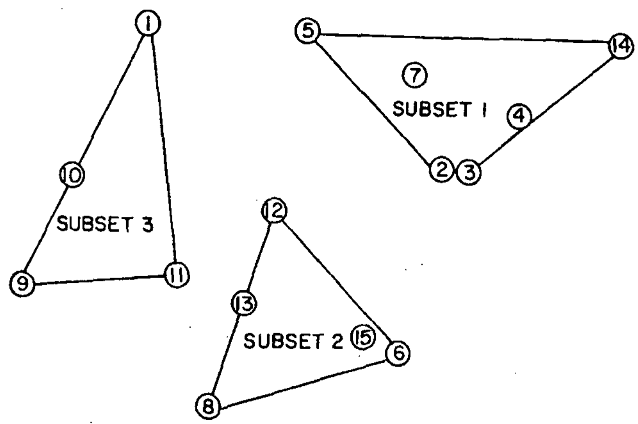
In multidimensional similarities scaling a high similarity score (i.e., a high interrelatedness number in a cell in Figures B.1 through B.4) means the points identifying the tasks connected by that score will be located close together in the multidimensional scaling space. Conversely, a low score means the points identifying the tasks connected by that score will be located further away in that space.

The two half matrices shown in Figures B.1 and B.2 constituted the data scaled by KYST in ascertaining the locations of the Level A tasks in two-dimensional space. The two matrices were scaled together, being treated as replicates to each other. The scaling procedure was carried out four times: in 4, 3, 2, and 1 dimensions, and the adequacy of the scaling solution was ascertained in each case. The two-dimensional solution was used to classify Level A tasks into subsets. This solution is presented in Figure B.5. In this figure, each numbered point marks the location of the task indexed by that number in a two-dimensional "task-interrelatedness" space, or map. The axes of the map are not named as they are arbitrary--what matters is the relative distance between the points in the space. Greater distance indicates less interrelatedness.

The two half matrices shown in Figures B.3 and B.4 constitute the data scaled by KYST in ascertaining the locations of the Level B tasks in two dimensional space. While the scaling was carried out in 4, 3, 2, and 1 dimensions, a satisfactory result was also obtained in two-dimensions. This solution is presented in Figure B.6, which may be interpreted in a manner similar to Figure B.5.

B.2.3.4 Fourth Stage - Grouping of Tasks Into Subsets

At Level A, the task interrelatedness map (Figure B.5) indicated that three distinct clusters of points could easily be identified. These comprised Tasks 2, 3, 4, 5, 7, 14 (Subset 1); Tasks 6, 8, 12, 13, 15 (Subset 2); and Tasks 1, 9, 10, 11 (Subset 3). These clusters met the criterion that there was a reasonable level of interrelatedness (i.e., closeness between points). Lines have been drawn in Figure B.5 to indicate the boundaries of the three clusters, and in each case the clusters occupy only a small area of the total map-space, and are distinct from one another. The tasks grouped within each cluster should be reasonably homogeneous with respect to the PSFs subject matter experts considered important in assessing the likelihood of successful performance.



The numbers refer to Level A tasks. The lines connecting the numbers delineate the three subsets of Level A tasks.

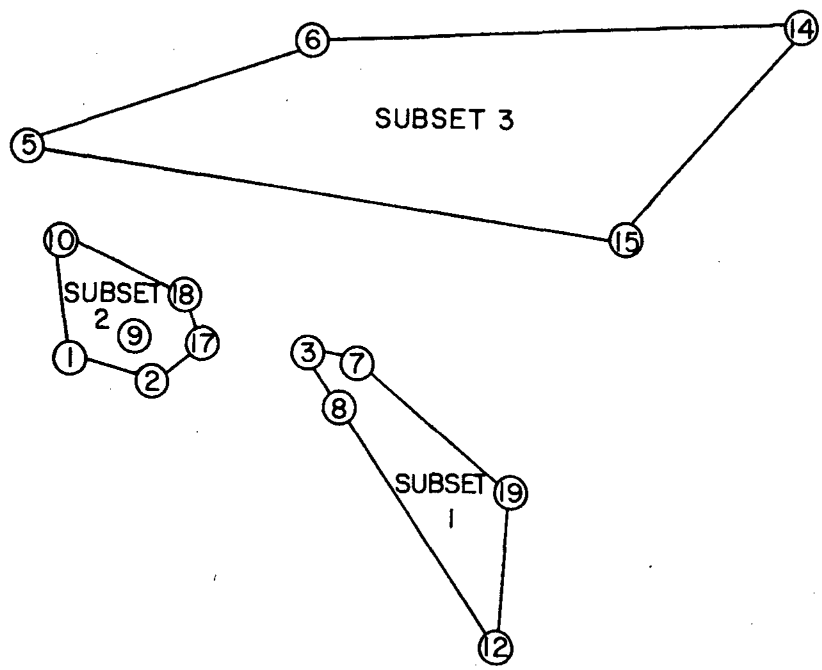
Figure B.5 Two-dimensional similarities scaling map of Level A task interrelatedness.

At level B, the task interrelatedness maps (Figure B.6) also indicated that three distinct clusters of points could be identified. These comprised Tasks 3, 7, 8, 12, 19 (Subset 1); Tasks 1, 2, 9, 10, 17, 18 (Subset 2); and Tasks 5, 6, 14, 15 (Subset 3). Subsets 1 and 2 formed tight and clearly distinct clusters, but in Subset 3 the four tasks formed a more diffuse cluster; however, it was distinct from the other two clusters. This solution was considered adequate as a basis for identifying three subsets of Level B tasks to be assessed together in SLIM-MAUD.

Summary of Task Classification Procedures

To summarize, Level A and B tasks were classified into subsets for the SLIM-MAUD test evaluation. Subject matter experts clarified task descriptions and then used paired comparison techniques to make judgments of the similarity of PSFs between tasks. Group consensus was achieved through discussions. These judgments of task similarity or interrelatedness were analyzed using a multidimensional similarities scaling procedure which resulted in the formation of three subsets of Level A tasks and three subsets of Level B tasks. These subsets were:

Level A tasks:	Subset 1:	Tasks 2, 3, 4, 5, 7, 14
	Subset 2:	Tasks 6, 8, 12, 13, 15
	Subset 3:	Tasks 1, 9, 10, 11



The numbers refer to Level B tasks. The lines connecting the numbers delineate the three subsets of Level B tasks developed.

Figure B.6 Two-dimensional similarities scaling map of Level B task interrelatedness.

Level B tasks:	Subset 1:	Tasks 3, 7, 8, 12, 19
	Subset 2:	Tasks 1, 2, 9, 10, 17, 18
	Subset 3:	Tasks 5, 6, 14, 15

B.2.4 Stage 3: Selection of the Members of the Four Subject Matter Expert Groups for Stage 4

To fulfill the need for an appropriate range of expertise within any group of subject matter experts, each group of experts selected to participate in the SLIM-MAUD test comprised participants who together possessed expertise in these three areas:

- Human Factors
- Probabilistic Risk Assessment
- Plant Operations

The precise composition of each group and the venue at which it met were as follows:

Group 1: One Human Factors specialist, two Probabilistic Risk Assessment experts, one expert with Operations experience; meeting at Brookhaven National Laboratory.

Group 2: One Human Factors specialist, two Probabilistic Risk Assessment experts, one expert with Operations experience; meeting at Brookhaven National Laboratory.

Group 3: One Human Factors specialist, one Probabilistic Risk Assessment experts, and two experts with Operations and Operator training experience; meeting at the NRC in Maryland.

Group 4: Two Human Factors specialists, one Probabilistic Risk Assessment experts, and one expert with Operations and Operator training experience; meeting at the NRC in Maryland.

B.2.5 Stage 4: Use of SLIM-MAUD by each Subject Matter Expert Group for each Subset of Tasks, Followed by Direct Numerical Assessment for all Tasks by each Group Member

Each of the four groups of subject matter experts met for one day and used the SLIM-MAUD software to assess the three Level A subsets of tasks and the three Level B subsets of tasks. Each task subset constituted one SLIM-MAUD session, thus, there were six SLIM-MAUD sessions for each group.

One member of each group was nominated to type in the group's responses at the keyboard during each SLIM-MAUD session. A facilitator and a technical recorder were also present during these sessions.

The facilitator had no previous experience with conducting SLIM-MAUD sessions. He was given the material in Attachment B-1, Section 3, "SLIM-MAUD Facilitator's Guide" which explained his role and provided information about conducting group sessions. The facilitator's role was designed to be minimal. His major role was to ensure that the viepoints of all group members were fully considered, and to guide the group in reaching consensus during each part of the SLIM-MAUD assessment process.

The technical recorder was a passive participant in the SLIM-MAUD sessions. His role was to observe and record the group discussions and to note the major points raised by group members during discussions.

Upon arrival, each group member was given a booklet containing task descriptions arranged by subset, and was asked to familiarize himself or herself with the task descriptions. Then the facilitator gave the group a brief, nontechnical introduction to SLIM-MAUD which appears in Attachment B-1, Section 4.

Following this introduction, questions from the group were answered and the group commenced the first SLIM-MAUD session. From here on, the facilitator's intervention in the group process was minimal; all the steps in the assessment procedure were controlled by MAUD in direct interaction with the group. The facilitator's few interventions were almost exclusively concerned with ensuring that the viewpoints of all the group members were fully considered in forming each judgment input to MAUD. In this way, consensus (or very occassionally an agreed compromise) was reached by the group on all aspects of the (decomposed) assessment of the tasks. In no case did a group member withdraw from the judgment process or indicate that his or her viewpoints were not represented in the interactions with MAUD.

SLIM-MAUD summary reports for each group of experts are reproduced in Attachment B-2. Attachment B-2 also contains the technical recorder's notes of the group discussions during the interactions with MAUD. Most of this discussion involved elaboration of task statements and PSFs.

A SLIM-MAUD summary report was printed for each subset of tasks assessed, resulting in six summary reports per group of subject matter experts. These summary reports are reprinted in Attachment B-2.

Immediately upon completion of the MAUD sessions, the four members in each group were assigned to separate rooms and asked to complete a direct estimate response booklet following the format described in Comer et al. (1984, Section 3.2.1). The tasks assessed within this booklet were the same 15 Level A and Level B tasks as in the SLIM-MAUD sessions but were arranged in numerical order rather than subset order. These materials are in Attachment B-1, Section 5.

Next, each expert participant in each group completed a questionnaire designed to assess the acceptability of SLIM-MAUD to himself or herself, and to other potential users. This questionnaire is presented in Attachment B-1, Section 6.

B.2.6 Stage 5: Analysis and Interpretation of Results from SLIM-MAUD Sessions

The test application of SLIM-MAUD allowed us to evaluate the practicality, acceptability, and usefulness of the methodology. The results of these analyses with respect to the criteria and issues contained in Table B-1 are described in the following sections.

B.2.6.1 Practicality

For the purposes of this evaluation, practicality was defined by the issues of cost, subject matter experts, support requirements, transportability, expandibility, time requirements, ability to interface with the Human Reliability Data Bank, and implementability of the SLIM-MAUD procedures (issues P1-P8 in Table B.1). This test evaluation demonstrated that SLIM-MAUD fulfills the criterion of practicality--it is relatively inexpensive in terms of costs, equipment, personnel, and time requirements, requires a minimum of training, is easy to implement in different locations, is applicable to a wide range of tasks, and is compatible with the Human Reliability Data Bank (Comer et al., 1983). Each of the issues that constitute the criterion of practicality will be briefly discussed.

P1 - Cost. The fundamental cost for implementing SLIM-MAUD consists of two essential components--software and equipment. The proprietary SLIM-MAUD software programs can be obtained by purchasing a MAUD user's license for approximately 200 pounds sterling. (See Appendix A of Volume II for procedures for obtaining a MAUD End-user's License Agreement.) The program code for the non-proprietary SLIM-MAUD software (HEP.BAS) is contained in Appendix A. The minimum equipment required include a personal computer with a minimum of 64K random access memory (RAM) which runs under CP/M or IBM/PC DOS operating systems, two floppy disk drives, monitor, and printer. Other materials needed to implement SLIM-MAUD include floppy diskettes, the IBM system diskette, 5" x 8" index cards, photocopies of forms in Attachment A-2 and writing materials.

P2 - Subject Matter Experts. Four SLIM-MAUD users took part in each of the four test sessions constituting this evaluation, and this is the recommended number of users for implementing SLIM-MAUD. Each group should include at least one individual having nuclear power plant operating experience with the specific type of plant being assessed and one individual having human factors experience, as well as individuals with PRA experience. Although not recommended, it is possible to implement SLIM-MAUD with a two-member group if the areas of expertise represented include human factors experience and operating experience.

Additional analyses (described in Section B.2.6.4 of this appendix) suggested that for applications of SLIM-MAUD to complex tasks, such as Level A, groups should be well represented with members having plant operating experience. For applications to simple tasks, such as Level B, human factors experts should be well represented within groups.

P3 - Support Requirements. When the group of subject matter experts are inexperienced with SLIM-MAUD, it is useful to have a facilitator present to provide an introduction to the SLIM-MAUD procedures and to guide group inputs. When the group does include an individual familiar with SLIM-MAUD, that person can act as facilitator.

P4 - Transportability. SLIM-MAUD can be implemented in a wide variety of settings, provided that the requirements enumerated above (P1-P3) are available at each location. The test application of SLIM-MAUD was implemented in two different locations using separate personal computers. The fact that SLIM-MAUD may be implemented on a variety of personal computers using the two most common operating systems, and the availability of personal computers in general, adds to the portability of the procedure. Personal computers can be transported from one location to another easily, or rented for short periods of time.

P5 - Expandibility. The assumptions underlying SLIM-MAUD and the SLIM-MAUD procedures themselves are sufficiently robust to be capable of assessing virtually any human task in nuclear power plants, although it is particularly useful for assessing complex tasks like the Level A tasks used in this test application.

P6 - Time Requirements. Three factors determine the time required to conduct a SLIM-MAUD session: user experience, number of tasks, and task complexity. A group of experienced SLIM-MAUD users can assess approximately 25 complex tasks (such as the Level A tasks) and as many as 60 simple tasks (such as the Level B tasks) in one working day. In this test application of SLIM-MAUD, the average total time taken by inexperienced groups to assess 15 complex tasks and 15 simple tasks was less than six hours. Table B.2 shows the average time requirements for each activity during the SLIM-MAUD test application. No group deviated more than 25% from the average time shown.

P7 - Interface With Human Reliability Data Bank. Level A tasks used in this test application correspond to Level 1 tasks of the Human Reliability Data Bank (Comer, et al., 1983). Level B tasks in this application correspond to Levels 2 and 3 tasks of the Human Reliability Data Bank. The selection of these tasks for the SLIM-MAUD test application ensured compatible interface of the SLIM-MAUD methodology with industry-specific data.

P8 - Implementability of Procedure. SLIM-MAUD was successfully implemented by a facilitator who did not take part in the development of the methodology, and who had minimal training in its application. Because SLIM-MAUD software is interactive, all interactions with, and data input to SLIM-MAUD can be accomplished by user groups without previous training.

B.2.6.2 Acceptability

Acceptability of the SLIM-MAUD methodology was defined by its acceptance by the scientific community, expert participants, potential users, NRC, and nuclear utilities (issues A1-A5 in Table B.1). Each of these issues are discussed below.

A1 - Scientific Community

The SLIM-MAUD test application was carried out with sufficient rigor to produce results that meet the standards of publication in reputable scientific journals. Journal publication of the test application findings will appear soon after the publication of this report.

Table B.2 Average Time for Completion of Each Activity in the SLIM-MAUD Test Application

Activity	Time Taken (Minutes)
1. Individual participants familiarize themselves with Level A and B tasks	30
2. Introduction to SLIM-MAUD (given by group facilitator)	30
3. Assessment of Level A Subset 1 tasks	110
(Break for refreshments)	
4. Assessment of Level A Subset 2 tasks	85
5. Assessment of Level A Subset 3 tasks	50
6. Assessment of Level B Subset 1 tasks	20
7. Assessment of Level B Subset 2 tasks	15
8. Assessment of Level B Subset 3 tasks	15
(Break for refreshments)	
9. Direct rating of Level A and B tasks by experts (each working alone)	45
10. Completion of SLIM-MAUD acceptability	
Questionnaire by experts (each working alone)	15

A2 - Expert Participants

A questionnaire was administered to the 16 experts who participated in the four SLIM-MAUD evaluation sessions (see Attachment B-1, Section 6 of this appendix). The major purpose of the questionnaire was to assess the level of acceptability of the SLIM-MAUD methodology to the expert participants.

The questionnaire addressed three key features within the issue of SLIM-MAUD's acceptability to expert participants: the degree of ease or difficulty in using SLIM-MAUD, the ability of SLIM-MAUD to elicit and organize judgments, and the meaningfulness of the results produced. Eight statements were used to assess these three features. Participants responded by selecting a value on a five-point Likert scale (strongly agree to strongly disagree.) Table B.3 summarizes the responses to these eight statements.

Overall, a majority of participants either strongly agreed or agreed with each of the eight statements. Participants expressed quite favorable evaluations of SLIM-MAUD, thus indicating the acceptability of the methodology.

For each of the three features of acceptability, the detailed results are as follows:

Table B.3 Results from Questionnaire Evaluating SLIM-MAUD to Expert Participants.

	Percent of Respondents Who				
	Strongly Agree	Agree	Neutral	Dis-agree	Strongly Disagree
<u>Ease or Difficulty of Using SLIM-MAUD:</u>					
- The SLIM-MAUD procedures were easy to understand.	19 (N=3)	75 (N=12)	6 (N=1)	0 (N=0)	0 (N=0)
- The SLIM-MAUD procedures could be clearly followed.	31 (N=5)	50 (N=8)	19 (N=3)	0 (N=0)	0 (N=0)
<u>Ability to Elicit and Organize Judgment:</u>					
- The SLIM-MAUD procedures effectively used my knowledge and experience.	12 (N=2)	69 (N=11)	19 (N=3)	0 (N=0)	0 (N=0)
- The SLIM-MAUD procedures seemed orderly.	31 (N=5)	44 (N=7)	25 (N=4)	0 (N=0)	0 (N=0)
- Using SLIM-MAUD improved my understanding of the factors that influence the likelihood of an incorrect action.	6 (N=1)	50 (N=8)	25 (N=4)	0 (N=3)	0 (N=0)
<u>Meaningfulness of Results Produced:</u>					
- The SLIM-MAUD procedures led to results that seemed meaningful.	19 (N=3)	63 (N=10)	12 (N=2)	6 (N=1)	0 (N=0)
- It would be easy to determine which PSFs had the greatest impact on HEP estimates by reviewing the SLIM-MAUD log of results.	12 (N=2)	63 (N=10)	19 (N=3)	6 (N=1)	0 (N=0)
- SLIM-MAUD can be useful to HRA segments of PRAs.	12 (N=2)	73 (N=12)	6 (N=1)	6 (N=1)	0 (N=0)

Ease of Use. The vast majority of experts (94%) said the SLIM-MAUD procedures were easy to understand and a solid, but slightly lower, majority (81%) said the SLIM-MAUD procedures could be clearly followed. Conversely, virtually none of the participants answered "Disagree" or "Strongly Disagree" to either of the ease of use statements.

Organizing Judgments. A strong majority of participants indicated the SLIM-MAUD procedures effectively tapped their knowledge (81%), and were orderly (75%), while the remaining respondents expressed a neutral opinion on these items. SLIM-MAUD was perceived to be weakest in its ability to help with an understanding of the factors that influence the likelihood of human error. This question received the greatest number of disagree responses (19%) from among all the question examined. Nonetheless, a majority of the participants (56%) agreed that SLIM-MAUD did help them understand the factors affecting the chances of an operator error.

Meaningfulness of the Results and Face Validity. As before, a strong majority of participants expressed favor about the results produced by SLIM-MAUD: 82% said the results seemed meaningful, 75% thought that the PSFs having greatest influence on HEP estimates were easily traceable, and 87% thought SLIM-MAUD to be useful to HRA segments of PRA. Only one respondent disagreed with each of the three statements addressing the meaningfulness issue, with an additional one to three users expressing neutral positions.

In addition to the above questions, which were designed to assess the level of acceptability among the expert participants themselves, two additional questions were asked to determine whether the participants thought other potential users would find SLIM-MAUD an acceptable methodology. In particular, participants were asked to respond to the following two questions on a five-point Likert scale (very likely to very unlikely).

- What is the likelihood that SLIM-MAUD will be acceptable for use by government and research PRA experts?

- What is the likelihood that SLIM-MAUD will be acceptable for use by utility PRA experts?

To the first question, two of the participants (12%) said "very likely," 12 (75%) said "likely," one (6%) said "neither" and the final one (6%) said "unlikely." Thus, the participants were in general, but not unanimous, agreement that PRA researchers would find SLIM-MAUD acceptable.

There was less widespread agreement on the second question, though here a majority (69%) of participants felt SLIM-MAUD would be acceptable to utility PRA experts. In particular, three (19%) said it was "very likely," eight (50%) said it was "likely" and one said "neither." Another one said "unlikely" and the final three said "very unlikely."

Taken together, the results show fairly widespread support for the acceptability of SLIM-MAUD as a method for estimating HEPs. Participants found the method acceptable to them and, according to their perception, to other potential users as well.

B.2.6.3 Usefulness

With regard to assessing the usefulness of SLIM-MAUD as a human reliability assessment technique, Table B.1 indicated that the three major issues were: (1) reliability, (2) face validity, and (3) covergent validity of the technique. The reliability and validity analyses were performed on the raw data contained in Appendix C of this volume. The results of these analyses are discussed below. Note that all analyses were done separately for Level A and B tasks, since it was expected that the usefulness of subjective judgment techniques depended upon the complexity of the tasks.

U1 - Reliability

The reliability of SLIs produced via SLIM-MAUD sessions was investigated by calculating the consistency of SLIs produced by the four groups of subject matter experts for Level A and B tasks. A single SLI represented each group's assessment of the success likelihood for a task. Thus, reliability of SLIs refers to "inter-group" reliability.

To evaluate the reliability of SLIs relative to other techniques of subjective judgment, measures of inter-judge consistency for direct HEP estimates of Level A and B tasks made by the SLIM-MAUD subject matter experts were calculated. Next, measures of inter-judge consistency for the direct HEP estimates made by subject matter experts who participated in Comer et al.'s evaluation of psychological scaling techniques were calculated using the data reported in Comer et al. (1984, Tables B.3 and B.4). This latter group evaluated the same tasks used in the SLIM-MAUD test application.

The results of these reliability analyses are reported below.

1. Inter-group Reliability of SLI Scores

The SLI scores contained in Table C.1 of Appendix C were used to construct six matrices of correlations between groups' SLI scores within each subset of tasks within Level A and B. Then the three intercorrelation matrices at Level A were aggregated to produce an intercorrelation matrix of inter-group reliability coefficients using the following formula:

$$r(agg) = \sqrt{\frac{\sum n_i r_i |r_i|}{\sum n_i}} \qquad (1)$$

where:

r_i = intercorrelation between two groups' scores for subset i of tasks
n_i = number of tasks in subset i (i = 1,2,3 at each level)
$|r_i|$ = absolute value of r_i.

The same was then done for the three intercorrelation matrices at Level B. This resulted in two intercorrelation matrices of inter-group reliability coefficients for Level A and B tasks which are presented in Table B.4.

Table B.4 Intercorrelations of SLI Values Between Groups Aggregated Over Subsets at Each Level

LEVEL A TASKS				
Group 1	-			
Group 2	.52	-		
Group 3	.73	.68	-	
Group 4	.61	.47	.69	-
	GR1	GR2	GR3	GR4

Overall inter-group correlation (r_{rms}) = .62

LEVEL B TASKS				
Group 1	-			
Group 2	.68	-		
Group 3	.60	.68	-	
Group 4	.77	.52	.65	-
	GR1	GR2	GR3	GR4

Overall inter-group correlation (r_{rms}) = .65

The overall inter-group reliabilities among all four groups was found by computing the root mean square of all the coefficients in these matrices, using the formula:

$$r(rms) = \sqrt{\sum_{i=1}^{n} \frac{r|r|}{n}} \quad (2)$$

where:

n = total number of entries in lower half matrix (excluding main diagonal) = 6

Using formula (2), the coefficients of overall inter-group reliability of SLI values were:

<u>LEVEL A</u> r(rms) = 0.62
(df = 50, $p < 0.01$)
range = +0.47 to +0.73

<u>Level B</u> r(rms) = 0.65
(df = 50, $p < 0.01$)
range = +0.52 to +0.77

2. Inter-group Reliability of HEP Estimates from SLIM-MAUD Subject Matter Experts Using Direct Psychological Scaling Techniques

After completing all SLIM-MAUD sessions, each subject matter expert made direct numerical estimates of HEPs and uncertainty bounds for each task following the format in Comer et al. (1984, Volume II, Appendix A, Section 3.2.1). This allowed us to compare the inter-group reliability coefficients of estimates generated by alternative subject judgment techniques (i.e., SLIM-MAUD and direct psychological scaling) when the same groups of experts used both techniques.

The direct HEP assessments made by the four experts within each group for each Level A and B tasks were aggregated using the formula proposed by Comer et al. (1984, p. A-27), viz:

$$\text{aggregate HEP} = \frac{\text{antilog } (\sum_{i=1}^{m} \log HEP_i)/m}{\text{antilog } [\sum_{i=1}^{m} \log (1-HEP_i)]/m + \text{antilog } (\sum_{i=1}^{m} \log HEP_i)/m} \quad (3)$$

where:

m = number of experts in each group = 4.

The resulting aggregated HEPs are presented in Table C.3 in Appendix C. These HEPs were converted to the logarithms of the HEPs before inter-group reliability coefficients were computed to avoid the potential biasing effects of using probabilities scaled from 0 to 1 which would introduce nonlinearity into the reliability correlation technique. Log HEPs are linearly related to SLI scores, and they have a constant unit of measurement along the entire length of the scale described and used by Comer et al. (1984).

Two matrices of intercorrelations for aggregated log HEPs were computed between groups across the 15 level A and 15 Level B tasks. These are shown in Table B.5.

Table B.5 Intercorrelations of Aggregated Direct Estimates of Log HEPs Level B tasks between Groups

LEVEL A TASKS				
Group 1	-			
Group 2	-.14	-		
Group 3	.22	.42	-	
Group 4	-.07	.65	.39	-
	GR1	GR2	GR3	GR4

LEVEL B TASKS				
Group 1	-			
Group 2	.42	-		
Group 3	.82	.63	-	
Group 4	.48	.32	.51	-
	GR1	GR2	GR3	GR4

The overall inter-group reliability exhibited in each of these two intercorrelation matrices was found by computing the root mean square of all correlations in each matrix, using formula (2) above. This gave the following results.

Level A	r(rms) = 0.36 (df = 0.65, p < 0.01) range = -0.14 to +0.65	Level B	r(rms) = 0.55 (df = 65, p < 0.01) range = +0.32 to +0.82

3. Inter-Judge Reliability of HEP Estimates from Independent Subject Matter Experts using Direct Psychological Scaling Techniques

Reliability analyses were also conducted using the direct numerical estimates of HEPs generated by the subject matter experts who participated in the Comer et al. (1984) evaluation of psychological scaling methods. These experts evaluated the same tasks evaluated during the SLIM-MAUD sessions. Thus, parallel reliability analyses on these data allowed us to compare the inter-judge consistency coefficients generated by alternative techniques (i.e., SLIM-MAUD and direct psychological scaling) for the same tasks evaluated by different experts.

These analyses are based on HEP estimates reported by Comer et al. (1984, Tables B-3 and B-4). Estimates for Level B tasks 4, 11, 13, 16, 20 were excluded since these tasks were excluded from the SLIM-MAUD test application.

Again, the HEP estimates were converted to log HEPs to avoid introducing bias into the reliability analyses. Figures B.7 and B.8 present the inter-judge reliability coefficients for log HEPs for Level A and B tasks, respectively.

The coefficients of overall inter-judge reliability were found by computing the root mean square of all the coefficients in each matrix, using formula (2), above. This gave the following results:

Level A	r(rms) = 0.43 (p < 0.01) range = -0.40 to +0.80	Level B	r(rms) = 0.63 (p < 0.01) range = -0.04 to +0.89

4. Discussion of Results of Reliability Analyses

Table B.6 summarizes the overall reliability coefficients of subjective estimates for Level A and B tasks generated by three methods: (1) SLIM-MAUD, (2) direct psychological scaling with the SLIM-MAUD subject matter experts, and (3) direct psychological scaling using a different group of subject matter experts.

Judge

Judge	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	-																		
2	.29	-																	
3	.36	.28	-																
4	.37	.20	.70	-															
5	.67	.37	.46	.45	-														
6	.41	.16	.55	.59	.54	-													
7	.65	.07	.62	.77	.61	.71	-												
8	.34	.32	.39	.15	.29	.49	.43	-											
9	.45	-.16	-.08	.09	.44	-.22	.14	-.18	-										
10	.57	.24	.56	.60	.60	.35	.50	-.04	.50	-									
11	.43	.59	.49	.50	.52	.74	.49	.28	-.40	.33	-								
12	.46	.01	.18	.57	.50	.54	.50	-.06	.33	.63	.40	-							
13	.54	.20	.41	.50	.22	.23	.47	.38	.06	.20	.26	.37	-						
14	.48	.26	.16	.43	.58	.47	.55	.36	.05	.22	.57	.49	.34	-					
15	.66	.06	.22	.31	.48	.48	.54	.67	.19	.16	.31	.33	.58	.69	-				
16	.35	.30	.16	.15	.28	.47	.21	.41	-.04	.45	.46	.30	-.05	.49	.48	-			
17	.71	.31	.33	.65	.64	.67	.80	.32	.23	.62	.60	.72	.44	.60	.55	.47	-		
18	.54	.35	.04	.26	.37	.26	.36	.16	.09	.40	.48	.54	.42	.75	.52	.60	.62	-	
19	.23	.27	.14	.04	.61	.24	.01	-.06	.14	.40	.44	.36	-.05	.50	.23	.45	.19	.54	-

Judge

Figure B.7 Intercorrelations between direct estimates of Log HEPs for 15 Level A tasks computed on data from Comer et al. (1984, Table B.3).

Judge	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	-																		
2	.62	-																	
3	.71	.63	-																
4	.48	.57	.47	-															
5	.67	.43	.80	.51	-														
6	.54	.69	.72	.62	.80	-													
7	.71	.34	.49	.56	.67	.60	-												
8	.69	.75	.87	.56	.75	.85	.51	-											
9	.70	.56	.66	.53	.69	.72	.78	.58	-										
10	.72	.88	.67	.54	.52	.73	.45	.88	.50	-									
11	.64	.85	.72	.38	.46	.57	.31	.76	.54	.86	-								
12	.50	.26	.77	.31	.56	.42	.22	.51	.44	.30	.31	-							
13	.80	.51	.89	.64	.92	.75	.70	.76	.80	.58	.54	.75	-						
14	.57	.52	.73	.44	.77	.89	.48	.85	.56	.70	.50	.51	.72	-					
15	.72	.62	.70	.54	.74	.79	.69	.75	.70	.74	.55	.46	.77	.82	-				
16	.71	.85	.83	.50	.69	.83	.43	.84	.70	.80	.74	.55	.75	.81	.80	-			
17	.57	.40	.69	.50	.78	.81	.60	.64	.80	.48	.42	.51	.80	.83	.80	.73	-		
18	.43	.40	.22	.47	.18	.07	.07	.24	-.04	.45	.27	.31	.32	.07	.31	.23	-.03	-	
19	.81	.51	.78	.46	.72	.68	.53	.76	.66	.61	.47	.78	.84	.76	.68	.77	.65	.27	-

Judge

Figure B.8 Intercorrelations between direct estimates of Log HEPs for 15 Level B tasks computed on data from Comer et al. (1984, Table B.4).

Table B.6 Overall Reliability Coefficients for Human Reliability Assessments of Level A and B Tasks Using SLIM-MAUD and Two Applications of Psychological Scaling Techniques.*

Technique	Task Level	
	A	B
SLIM-MAUD assessments	0.62	0.65
Direct HEP estimates using psychological scaling, SLIM-MAUD subject matter experts	0.36	0.55
Direct HEP estimates using psychological scaling, Comer et al. (1984) experts	0.43	0.63

*$p < .01$ for all correlation coefficients.

SLI assessments were more reliable than direct numerical estimates of HEPs for Level A tasks when the different techniques were applied by the same experts (+0.62 vs +0.36) and when the different techniques were applied by different experts (+0.62 vs +0.43). The three techniques did not appear to differ in reliability when applied to less complex (i.e., Level B) tasks. Direct estimation techniques appeared to produce more reliable results when applied to less complex tasks. However, SLIM-MAUD appeared to produce equally reliable results for assessing both complex and simple tasks (i.e., +0.62 and +0.65, respectively).

A similar pattern of results was revealed when the respective ranges of the reliability coefficients for each technique was examined (see Tables B.4 and B.5). The range of between-group reliability coefficients for SLIM-MAUD for Level A was +0.47 to +0.73, and for Level B the range was +0.52 to +0.77. This indicates that there was always at least moderate agreement between any two groups of experts, no matter which pair of groups was compared. In contrast, the ranges of the between-group and between-judge reliability coefficients for Level A task assessments using the direct psychological scaling methods were very large (-0.14 to +0.65 and -0.40 to +0.80).

We may conclude that SLIM-MAUD is a more reliable procedure than direct psychological scaling for estimating HEPs, and that this advantage is particularly marked when assessing Level A tasks.

The direct numerical estimates of HEPs produced by the subject matter experts who participated in the Comer et al. (1984) study appeared to be somewhat more reliable than the direct numerical estimates of HEPs produced by the subject matter experts who participated in the SLIM-MAUD evaluation. This held true for both Level A tasks: +0.43 vs +0.36, respectively, and Level B tasks: +0.63 vs +0.55, respectively. This may have occurred because the experts who participated in the Comer et al. study were very homogeneous in terms of areas of expertise. All 19 subject matter experts were NRC-certified BWR instructors. In contrast, the subject matter experts who participated in the SLIM-MAUD evaluation represented expertise in the areas of human factors, operations, and PRA.

U2 - Face Validity of SLIM-MAUD

Face validity refers to whether the procedures appear relevant, appropriate, and valid to users of the methodology. It is considered an essential precursor to more rigorous types of validity, and is a desirable feature of any methodology. Face validity provides some assurance that a methodology measures what it is supposed to be measuring.

The face validity of SLIM-MAUD was assessed with three items on the questionnaire administered to the session participants at the end of the SLIM-MAUD sessions. Subject matter experts were asked to respond on a five-point Likert scale (strongly agree to strongly disagree). Results obtained for the three items which refer to face validity are presented in Table B.7

A strong majority of participants made favorable evaluations of the face validity of the SLIM-MAUD methodology: 82% said the results seemed meaningful; 75% thought that the PSFs having the greatest influence on HEP estimates were easily traceable; and 87% thought SLIM-MAUD to be used to HRA segments of PRA. One respondent disagreed with each of the three statements addressing the meaningfulness issue, and three respondents expressed a neutral position to one or two face validity items. Thus, the evidence in general shows SLIM-MAUD to have an acceptable level of face validity.

U3 - Convergent Validity of SLI Values with Estimates Provided by Other Tecnhiques

It is difficult to validate the utility of subjective judgment techniques for estimating HEPs in nuclear power plant operations because of the low frequency of recorded operator failures, especially for complex tasks. Thus, there is no objective criteria against which to compare the subjective probability estimates. Therefore, the criterion-related validity of the SLIM-MAUD methodology was assessed by observing its "convergence" with other methods for estimating HEPs in nuclear power plants.

Convergent validity refers to the size of the correlation between methodologies which are designed to measure the same construct. Thus, if the SLIM-MAUD methodology produced valid assessments, we would expect the SLI results to correlate highly with the results of other reliable subjective judgment techniques applied to the same tasks.

B.7 Results of Questionnaire Items Assessing Face Validity of SLIM-MAUD

	Percentage of Respondents Who:				
	Strongly Agree	Agree	Neutral	Disagree	Strongly Disgree
The SLIM-MAUD procedures led to results that seemed meaningful	19 (N=3)	63 (N=10)	12 (N=2)	6 (N=1)	0 (N=0)
It would be easy to determine which PSFs had the greatest impact on HEP estimates by reviewing the SLIM-MAUD log of results	12 (N=2)	63 (N=10)	19 (N=3)	6 (N=1)	0 (N=0)
SLIM-MAUD can be useful to HRA segments of PRAs	12 (N=2)	75 (N=12)	6 (N=1)	6 (N=1)	0 (N=0)

The degree of convergence between the results of two methodologies will be influenced by the reliability of each of the methodologies. The inter-group reliability coefficients for the results of SLIM-MAUD and direct psychological scaling of HEPs indicated that the methodologies possessed different levels of reliability (see Table B.6), ranging from low to moderate. This measurement error will tend to decrease the size of the correlation coefficient between SLIM-MAUD and any other methodology. However, to the extent that SLIM-MAUD is capable of producing reasonably stable and appropriate estimates of HEPs, it should demonstrate acceptable levels of convergent validity with a range of similar methods.

The SLI assessments were correlated with direct HEP estimates from the Level A and B tasks produced by psychological scaling techniques from two groups of experts: (1) the 16 subject matter experts who participated in the SLIM-MAUD test application and (2) the 19 subject matter experts who participated in Comer et al.'s evaluation of psychological scaling techniques. In addition, the SLI assessments for the 15 Level A tasks were correlated with HEP estimates for these tasks given in the Handbook of Human Reliability Analysis (Swain and Guttmann, 1983).

1. Convergent Validity Coefficients between SLI Assessments and Direct HEP Estimates using Psychological Scaling by SLIM-MAUD Subject Matter Experts

SLI assessments were correlated with the direct HEP estimates produced by SLIM-MAUD subject matter experts using the data in Table C.1 in Appendix C (SLI scores) and Table C.3 in Appendix C (aggregated HEPs). This resulted in three intercorrelation matrices between SLI scores and aggregated HEPs at each task level. These intercorrelation matrices were then aggregated using formula 1 (page B-18) to produce two overall intercorrelation matrices for assessments of Level A and B tasks. These matrices are shown in Tables B.8 and B.9, respectively.

Table B.8 Intercorrelations Between the Four Subject Matter Expert Groups' SLI Scores and Their Aggregated Direct Assessments of Log HEPs for Level A Tasks.

	Group			
Group	1	2	3	4
1	0.67	0.76	0.74	-0.48
2	-0.08	0.23	0.35	-0.49
3	0.37	0.70	0.60	-0.35
4	0.50	0.37	0.74	0.40

Table B.9 Intercorrelations Between the Four Subject Matter Expert Groups' SLI Scores and Their Aggregated Direct Assessments of Log HEPs for Level B Tasks.

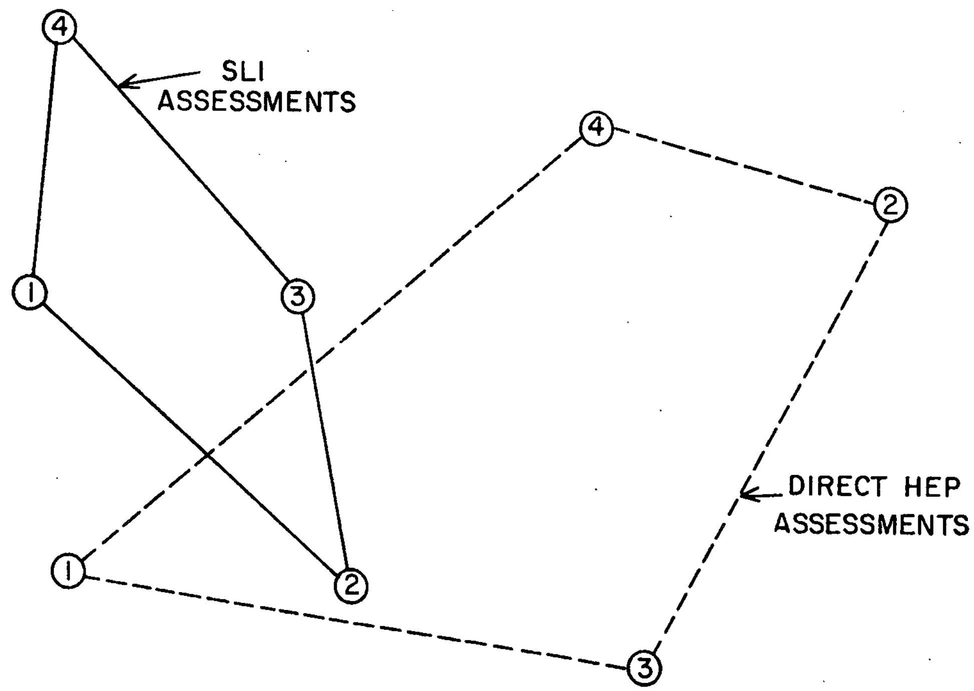
	Group			
Group	1	2	3	4
1	.90	.72	.56	.62
2	.82	.39	.64	.63
3	.84	.67	.45	.63
4	.57	.48	.29	.54

Each matrix gives the intercorrelations between group SLI scores and direct HEP estimates. These correlations are measures of the convergent validity of SLIM-MAUD results and direct psychological scaling results. The root mean squares of the correlations in the matrices shown in Tables B.8 and B.9 were computed as an index of the convergent validity of these methods for Level A and B tasks, respectively. These inter-method correlations were:

Level A: r(rms) = +0.48
Level B: r(rms) = +0.66

Multidimensional Scaling Analyses

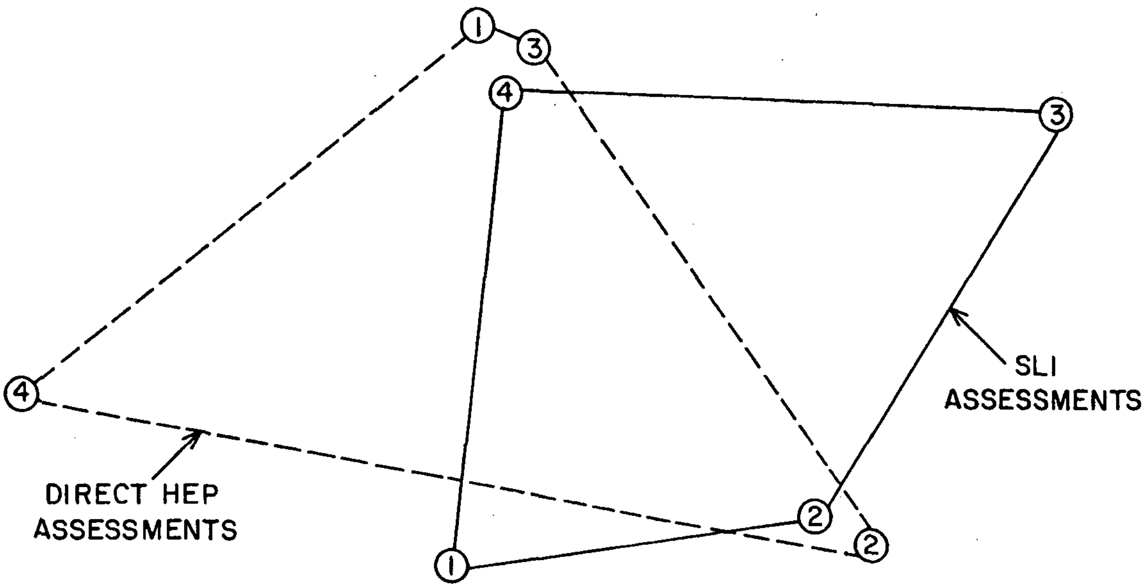
Multidimensional scaling analyses, using the computer program KYST, were carried out to provide a more detailed picture of the similarities and differences in the assessments made by the subject matter experts using SLIM-MAUD and direct psychological scaling of HEPs. Nonmetric similarity scaling analyses were performed on the data contained in Tables B.8 and B.9 in 4, 3, 2, and 1 dimensions, with satisfactory stress being obtained for the two-dimensional solutions. The resulting multidimensional scaling maps for Level A and B tasks are shown in Figures B.9 and B.10, respectively.



The numbers at each point of the two quadrilaterals represent the position in space for each subject matter expert group (Groups 1 to 4). SLI assessments are connected by a solid line. Direct numerical assessments are connected by a dotted line.

Figure B.9 Two-dimensional multidimensional scaling map of degree of intercorrelations between assessments of Level A tasks made by the 4 subject matter expert groups using SLIM-MAUD and psychological scaling.

The amount of overlap in each figure indicates the convergence between human reliability assessments made using SLIM-MAUD and direct psychological scaling of HEPs. For Level A tasks (Figure B.9), there was very little overlap between the two envelopes formed by SLI assessments and direct HEP estimates. For Level B tasks (Figure B.10), there is more overlap between the two types of assessments.



The numbers at each point of the two quadrilaterals represent the position in space for each subject matter expert group (Groups 1 to 4). SLI assessments are connected by a solid line. Direct numerical assessments are connected by a dotted line.

Figure B.10 Two dimensional multidimensional scaling map of degree of intercorrelations between assessments of Level B tasks made by the 4 subject matter expert groups when using SLIM-MAUD and psychological scaling.

2. Convergent Validity Coefficients between SLI Assessments and Direct HEP Estimates using Psychological Scaling by Subject Matter Experts Participating in the Comer et al. (1984) Evaluation

SLI assessments were correlated with the direct numerical estimates of HEPs produced by the subject matter experts participating in Comer et al. (1984) evaluation, using data contained in Table C.1 in Appendix C (SLI scores) and Table C.3 in Appendix C (aggregated log HEPs). The aggregated log HEP estimates were then correlated with the SLI scores produced by each log MAUD expert group for each subset of Level A and B tasks. These correlations were then aggregated across the three subsets at each level using formula 1, resulting in the overall correlations shown in Table B.10.

Table B.10 Correlations between SLIs Produced by each SLIM-MAUD Group and Direct Numerical Estimates of HEPs Computed by Comer et al. (1984).

	Tasks	
SLIM-MAUD Group	Level A	Level B
Group 1	0.59	0.76
Group 2	0.35	0.72
Group 3	0.75	0.62
Group 4	0.24	0.64
Overall r(rms) =	0.52	0.69

In general, the convergent validity coefficients between SLI assessments and direct estimated HEPs are higher for the Level B tasks than Level A tasks.

3. Convergent Validity Coefficients between SLI Assessments and HEP Estimates for the 15 Level B Tasks from the Handbook of Human Reliability Analysis

Following Comer et al.'s example (1984, Table B-12), HEP estimates for Level B tasks were correlated with the SLI assessments and direct HEP estimates using psychological scaling made by SLIM-MAUD subject matter experts. The results are presented in Table B.11.

Discussion

The pattern of correlations in Table B.12 indicates adequate convergent validity of SLIM-MAUD results and results generated by other methods for estimating human reliability, given the low to moderate reliability of the methods. In general, higher validity coefficients were obtained for Level B tasks than for Level A tasks which probably reflects the higher reliability of the direct psychological scaling techniques for level B tasks.

The results of the multidimensional scaling analyses of assessments made by different groups of experts discussed above provide additional information regarding the convergence of SLIM-MAUD results with the results of direct psychological scaling methods. For Level A tasks, the SLIM-MAUD assessments were more consistent and occupied a smaller and separate domain in the "reliability space" mapped by the multidimensional scaling analysis, than that occupied by the psychological scaling assessments. For Level B tasks, the domain mapped by each technique was similar in size with a substantial amount of overlap. This indicated greater convergent validity between SLIM-MAUD assessments and direct psychological scaling assessments for Level B tasks than for Level A tasks. The relative lack of consistency of direct HEP estimates for Level A tasks may have reduced the convergence between the assessments. In addition, the direct scaling techniques appeared to yield more "off center" (i.e., method-idiosyncratic) results.

Table B.11 Correlations Between SLI Values and Direct HEPs Assessed by the Four Subject Matter Expert Groups in the SLIM-MAUD Test Plan and the HEP Estimates Given in the Handbook of Human Reliability Analysis (Level B Tasks Only)

Group		SLIM-MAUD Correlation between Handbook Estimates and SLI Values	PSYCHOLOGICAL SCALING Correlation between Handbook Estimates and Aggregated Direct HEP Assessments
1		+0.48	+0.38
2		+0.74	-0.03
3		+0.50	+0.37
4		+0.39	-0.03
Overall	r(rms) =	+0.54	+0.26

Table B.12 Correlations Between SLI Assessments and Other Techniques for Human Reliability Analysis

Source of Data for Comparison	Task Level A	Task Level B
Direct HEP estimates made by SLIM-MAUD subject matter experts	0.48	0.66
Direct HEP estimates made by Comer et al. subject matter experts	0.52	0.69
HEP estimates from Handbook of Human Reliability Analysis	-	0.54

p < .01 for all correlation coefficients

B.2.6.4 Additional Analyses

The reliability analyses of SLIM-MAUD indicated that SLI assessments have a moderate degree of reliability between groups of subject matter experts assessing the same tasks. Although SLIM-MAUD estimates were more reliable than the estimates derived from psychological scaling techniques, it would still be desirable to increase the reliability of SLIM-MAUD estimates by eliminating potential sources of bias in the methodology.

Three potential sources of systematic error in SLI scores were examined:

- biases in selection of PSFs by individual subject matter expert groups,
- bias due to the inappropriate selection of subject matter experts,
- bias in task evaluations by individual subject matter expert groups.

B.2.6.4.1 Investigation of possibility of Bias in Selection of Performance Shaping Factors

At the beginning of the SLIM-MAUD sessions, the group facilitator instructed subject matter experts to consider the effects of performance shaping factors (PSFs) on task performance. Several types of PSFs were listed in order to help the experts think about the types of factors which might exert significant influences on task performance. We were concerned as to whether the availability and salience of the PSFs on the list might prevent the experts from generating other important PSFs. According to Tversky and Kahneman (1983), the availability of the information given in a situation may bias the judgment of experts.

Subject matter expert groups were quite selective in the PSFs they considered for each task. The average number of PSFs considered by each group for Level A tasks was 4.0; the average number of PSFs considered for Level B tasks was 2.6. This raised the question of whether this selectivity was appropriate or whether it represented "PSF availability" bias of the type described by Tversky and Kahneman.

We investigated the possibility that the list of PSFs used by subject matter experts to select PSFs was a source of potential bias in SLIM-MAUD assessments. SLIM-MAUD assessments could be biased if the major source of reliability among SLI values was the degree of similarity of the PSFs considered. If such a bias existed, when groups used highly similar PSF sets to assess tasks, we would expect high reliability coefficients between SLI scores. When groups used dissimilar PSF sets to assess tasks, we would expect low reliability coefficients between SLI scores.

This hypothesis was tested using multidimensional scaling techniques which examined the similarities and differences between the sets of PSFs employed by the four subject matter expert groups in the SLIM-MAUD test plan.

Two independent judges, familiar with SLIM-MAUD and the tasks being assessed in the SLIM-MAUD test plan, made judgments concerning the degree of similarity between each set of PSFs employed by each group for each subset of tasks. Each judge made pairwise comparisons between sets of PSFs employed by a particular group when assessing a particular task subset. The following information (extracted from the SLIM-MAUD session summaries) about sets of PSFs was summarized on index cards to facilitate the comparisons:

1. PSFs included when computing SLIs (together with their weight and ideal points.
2. PSFs deleted in interaction with MAUD.
3. PSFs deleted a priori by the group during discussion.

1 -9

Two Totally Dissimilar Sets of PSFs	Two Totally Similar Sets of PSFs

After making each assessment individually, the two judges compared their results and discussed any differences until agreement on each comparison was obtained. The resulting assessments of the similarities between the sets of PSFs are shown in Table B.13.

If the major source of variability in SLI assessments was the similarity of the PSFs used during the assessment process, then the judgments of PSF similarity should correlate with the similarities (i.e., degree of positive intercorrelation) between SLI values produced by the SLIM-MAUD experts. Therefore, each PSF similarity rating was correlated with the corresponding SLI score in Tables B.8 and B.9. The resulting correlations were:

Level A: $r = +0.30$ ($n = 18$, not significant)
Level B: $r = +0.02$ ($n = 18$, not significant)

These nonsignificant correlations indicated that the inter-group reliability in SLI values for Level A and B tasks was not primarily a function of similarity in the group's selection of PSF sets. Thus, it was concluded that the SLIM-MAUD procedures for the selection of PSFs are probably not a source of systematic error in the assessment.

Nonmetric multidimensional scaling analysis was carried out on PSF similarity ratings using the computer program KYST. Scaling was carried out in 4, 3, 2, and 1 dimensions, with satisfactory stress being obtained for the two-dimensional solutions. Two "PSF similarity maps" of the 12 (4 groups x 3 task subsets) sets of PSFs used by the subject matter experts in assessing the Level A and B tasks were generated. The results of this analysis are shown in Figure B.11 for Level A task subsets and in Figure B.12 for Level B task subsets.

Table B.13 PSF Similarity Ratings for the Level A and B Task Subsets (Higher Score Reflects Greater Similarity)

LEVEL A

TASK SUBSET 1

1	-			
2	5	-		
3	8	8	-	
4	6	8	7	-
Group	1	2	3	4

TASK SUBSET 2

1	-			
2	2	-		
3	4	5	-	
4	7	5	6	-
Group	1	2	3	4

TASK SUBSET 3

1	-			
2	6	-		
3	8	8	-	
4	2	2	3	-
Group	1	2	3	4

LEVEL B

TASK SUBSET 1

1	-			
2	4	-		
3	3	5	-	
4	3	4	4	
Group	1	2	3	4

TASK SUBSET 2

1	-			
2	3	-		
3	6	8	-	
4	3	3	6	
Group	1	2	3	4

TASK SUBSET 3

1	-			
2	3	-		
3	5	4	-	
4	6	5	4	
Group	1	2	3	4

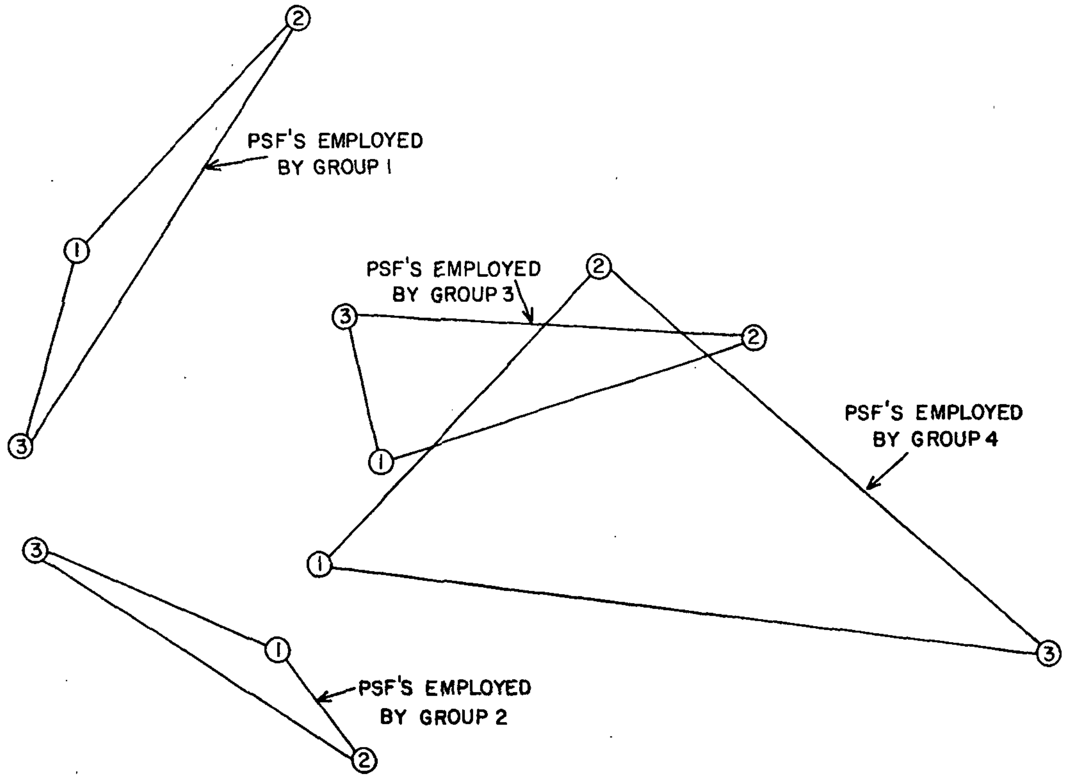


Figure B.11 Two-dimensional PSF similarity map for assessments of the three Level A task subsets made by the four groups of subject matter experts.

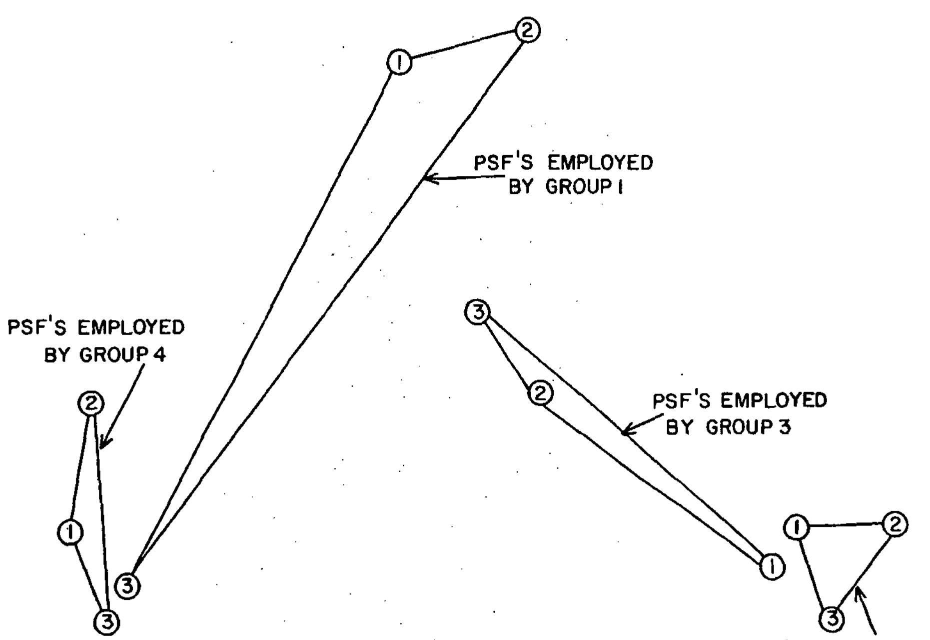


Figure B.12 Two-dimensional PSF similarity map for assessments of the three Level B task subsets made by the four groups of subject matter experts.

In each figure, the numbered points refer to each subset of tasks at Level A and B, and the respective set of PSFs employed by each group of subject matter experts for that subset. For example, in Figure B.7, the number 1 refers to the PSFs associated with subset 1 of Level A tasks, 2 refers to subset 2 of Level A, and so on. The three points identifying PSF sets used by each particular group are joined by lines, thus marking out the similarity envelope for the PSF sets used by that group.

Figures B.11 and B.12 show that for both Level A and B tasks, each group tended to use idiosyncratic PSF sets; that is, the envelopes defining the cluster of PSF sets are separated from each other in the PSF similarity space. This finding is particularly strong for Level B tasks, where there is no overlap among the sets of PSFs employed by each group.

Subject matter experts tended to select PSFs which were idiosyncratic to the group. The level and range of the group's expertise in human factors, probabilistic risk assessment, and nuclear power plant operations may have been important influences on the PSFs selected. Thus, the utility of SLI assessments may have been affected by the selection of the subject matter experts.

B.2.6.4.2 Investigation of the Possibility of Bias Due to Inappropriate Selection of Subject Matter Experts

Each subject matter expert group which participated in the SLIM-MAUD test application was composed so that the following areas of expertise were represented: (1) human factors, (2) probabilistic risk assessments, and (3) plant operations. In contrast, Comer et al.'s (1984) group of 19 BWR instructors all were experts in the area of plant operations. We attempted to investigate how the range of expertise represented by the group affected the assessments.

Due to constraints of the test plan, we were not able to systematically vary group composition to determine the effects on SLI assessments. However, the effects of major area of expertise on the inter-judge reliabilities of the direct estimates of HEPs for the Level A and B tasks were examined. The similarities and differences among these HEPs as a result of the subject matter experts' major areas of expertise were examine using nonmetric multidimensional scaling analyses.

The data used in these analyses were the inter-judge correlations between HEPs for each task scaled as log probabilities. These data are presented in Table B.14 for Level A tasks and Level B tasks and by type of expertise. Each judge is designated by a letter (A-P). Group 1 contained judges A-D, Group 2 contained judges E-H, Group 3 contained judged I-L, and Group 4 contained judged M-P.

Multidimensional scaling was carried out in 4, 3, 2, and 1 dimensions, with satisfactory stress being obtained for the two-dimensional solutions, resulting in the "HEP similarity maps" shown in Figures B.13 and B.14 for Level A and B tasks, respectively.

In each of these figures, the point referring to each expert is identified by a single letter code (A-P) code. The points identifying those experts with similar types of expertise are joined by a line delineating the "similarity envelope" for experts with that type of expertise. A continuous line is used to identify the envelope for operations expertise, a broken line for PRA expertise, and a dotted line for human factors expertise.

The area within the maps shown in Figures B.13 and B.14 represents the amount of similarity between HEP estimates for judges with similar types of experience. In Figure B.13, the line connecting subject matter experts with PRA expertise encloses a very large space, as do the lines connecting experts with human factors expertise. This suggests that in each case there was considerable divergence in the direct assessments of HEPs of Level A tasks between experts with each of these two types of expertise. In addition, since the area enclosed by the two types of expertise overlap, it is not known whether differences in expertise between PRA and human factors experts had any unique or consistent influence on their direct estimates of HEPs.

The lines connecting the experts with operations expertise enclose a very small space at the center of the plot. This suggests that experts with operations expertise showed less variability in their ratings of HEPs of the Level A tasks. Since the area representing experts with operations expertise is located at the center of the areas representing experts with PRA expertise and experts with human factors expertise, the average of their ratings is not different from the average of the HEP estimates made by experts with PRA or human factors expertise.

Figure B.14 shows that for Level B tasks, the lines connecting experts with human factors expertise enclose a much smaller space than the maps representing experts with operations expertise or PRA expertise. In addition, the map defining the human factors expertise does not overlap the maps defining the PRA and operations expertise clusters. This suggests that the domain of expertise represented by the human factors specialists was unique.

These results suggested that there were considerable differences in the reliability of the direct HEP estimates made by subject matter experts possessing different types of expertise. Experts with operations experience provided the most similar estimates of HEPs for Level A tasks and human factors specialists provided the most similar estimates of HEPs for Level B tasks.

In order to confirm these hypotheses, overall inter-expert reliability coefficients for direct HEP estimates for each type of expertise were calculated using the data from Table B.14. The overall inter-expert reliability coefficients for each of the six intercorrelations matrices shown in Table B.14 were found by calculating the root mean square of the coefficients in each matrix, using formula (2) from Section B.2.6.3. These results are presented in Table B.15.

Table B.14 Intercorrelations Between Individual Experts' Direct Estimates of Log HEPs for Level A and B Tasks, Grouped by Type of Expertise of the Expert

LEVEL A TASKS

(a) Human Factors Expertise

Expert	A	F	J	N	O
A	-				
F	-.29	-			
J	.13	-.07	-		
N	-.59	.42	.31	-	
O	-.12	.40	.00	.32	-

(b) Operations Expertise

Expert	D	G	I	L	P
D	-				
G	.95	-			
I	.86	-.01	-		
L	.58	.38	.69	-	
P	.61	.64	.62	.66	-

(c) PRA expertise

Expert	B	C	E	H	K	M
B	-					
G	.40	-				
E	-.26	-.22	-			
H	-.26	-.29	-.19	-		
K	.24	.11	-.03	-.02	-	
M	.52	.25	-.34	-.47	-.18	-

LEVEL B TASKS

(a) Human Factors Expertise

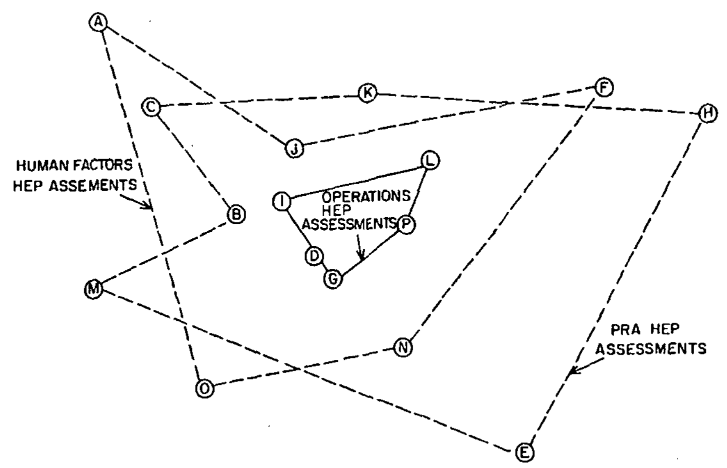
Expert	A	F	J	N	O
A	-				
F	.39	-			
J	.83	.65	-		
N	.74	.40	.88	-	
O	.67	.37	.63	.58	-

(b) Operations Expertise

Expert	D	G	I	L	P
D	-				
G	.78	-			
I	.20	.32	-		
L	.34	.48	.75	-	
P	-.07	.11	-.17	.15	

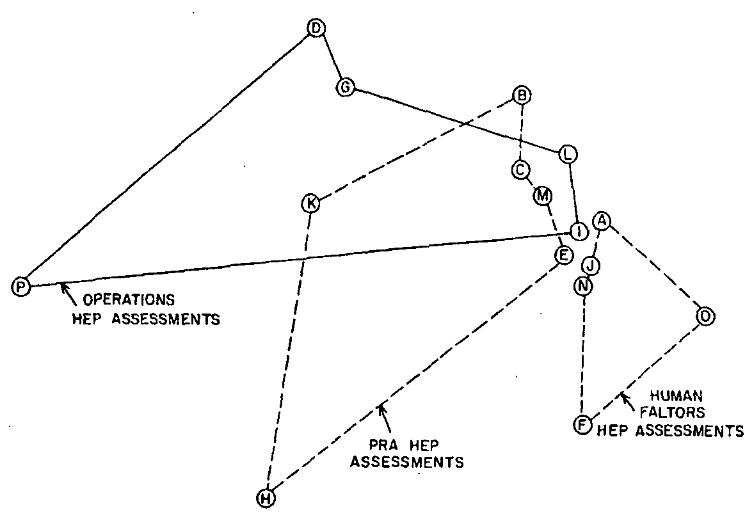
(c) PRA expertise

Expert	B	C	E	H	K	M
B	-					
C	.80	-				
E	.49	.70	-			
H	.00	-.07	.06	-		
K	.39	.55	.39	.23	-	
M	.64	.84	.76	-	.28	.36



- A continuous line connects assessments made by experts with operations expertise.
- A broken line connects assessments made by experts with PRA expertise.
- A dotted line connects assessments made by experts with Human Factors expertise.

Figure B.13 Two-dimensional inter-expert similarity map for direct estimates of HEPs for Level A tasks.



- A continuous line connects assessments made by experts with operations expertise.
- A broken line connects assessments made by experts with PRA expertise.
- A dotted line connects assessments made by experts with Human Factors expertise.

Figure B.14 Two-dimensional inter-expert similarity map for direct estimates of HEPs for Level B tasks.

Table B.15 Overall Reliability of Direct Estimates of HEPs Made by Experts Possessing Different Types of Expertise

Type of Expertise	Level A	Level B
Human Factors	+0.10	+0.64
Operations	+0.65	+0.41
Probabilistic Risk Assessment	-0.17	+0.50

Summary and Discussion

The multidimensional scaling maps revealed that for Level A tasks, experts with plant operations experience were more consistent (i.e., greater inter-judge agreement) in their HEP estimates than experts whose expertise lay in other areas. The overall inter-judge reliability coefficient was +0.65 for experts with plant operations expertise, +0.10 for experts with human factors expertise, and -0.17 for experts with PRA expertise. All HEP estimates for experts with plant operations expertise were located at the center of the inter-judge consistency map.

For Level B tasks, experts with human factors expertise were more consistent in their HEP estimates than experts whose expertise lay in other areas. The overall inter-judge reliability coefficient was +0.64 for human factors expertise, +0.50 for PRA expertise, and +0.41 for operations expertise. However, the "human factors cluster" of HEP estimates lay at one side of the inter-judge reliability map, with no overlap with the clusters representing judgments made by other experts. This indicates that the human factors experts had a consensus view about performance on Level B tasks which was quite different from the views held by the other experts, even though they had recently been exposed to these views in the group discussions during the SLIM-MAUD sessions.

On the basis of these analyses, the following recommendations are made for the composition of subject matter expert groups involved in human reliability assessment:

1. For Level A-type tasks, experts with operational expertise should be well represented; that is, at least two or three members of a four-person group. One or two group members with expertise in other areas is desirable to stimulate group discussion on the full range of issues to be considered (with human factors experts slightly preferred over PRA expertise in this respect). Experts with sources of expertise other than plant operations should not form a majority in a group assessing complex nuclear power plant operating tasks.

2. For Level B-type tasks, experts with human factors expertise should be represented since they appear to bring a unique perspective to assessing nuclear power plant performance. The group should consist of at least one, and preferably two persons with human factors experience, and one or two members with PRA and/or operations experience.

B.2.6.4.3 Investigation of the Possibility of Bias in the Elaboration of Task Definitions

SLIM-MAUD works best in concrete applications such as a PRA for a particular plant, or in generic applications where a specific plant is identified to typify that application. However, in the test application of SLIM-MAUD, task definitions did not include references to a specific plant. As a result, each subject matter expert group elaborated the definitions of many of the tasks in order to clarify the judgmental situation. For example, in one group there was much discussion among the subject matter experts concerning the task RCIC OPERATION AFTER LOSP (Level A, Task 2) which stated:

"During a Loss-of-offsite-power (LOSP) transient, the generator has tripped, the reactor has scrammed, and the normal feedwater system in inoperable. According to the procedures, the reactor water level should be recovered and maintained by manually operating the reactor core isolation cooling (RCIC) system. What is the likelihood that the operator will fail to operate the RCIc system correctly?"

The group felt that it was necessary to specify the type of BWR being assessed since time available to perform the task depends on the type of BWR. The group agreed to assume a BWR4.

Elaborations of task definitions differ among groups. Hence, an important source of variability in SLI assessments may be that after elaboration of the task definitions, different tasks were being assessed. Therefore, the reliability of SLI assessments might be considerably improved if care is taken to provide concrete and detailed task definitions which include the setting and environment. When possible, SLIM-MAUD sessions should be held within the plant being assessed in a specific PRA to make the task setting more concrete to the subject matter experts.

"During a Loss-of-offsite-power (LOSP) transient, the generator has tripped, the reactor has scrammed, and the normal feedwater system in inoperable. According to the procedures, the reactor water level should be recovered and maintained by manually operating the reactor core isolation cooling (RCIC) system. What is the likelihood that the operator will fail to operate the RCIc system correctly?"

The group felt that it was necessary to specify the type of BWR being assessed since time available to perform the task depends on the type of BWR. The group agreed to assume a BWR4.

Elaborations of task definitions differ among groups. Hence, an important source of variability in SLI assessments may be that after elaboration of the task definitions, different tasks were being assessed. Therefore, the reliability of SLI assessments might be considerably improved if care is taken to provide concrete and detailed task definitions which include the setting and environment. When possible, SLIM-MAUD sessions should be held within the plant being assessed in a specific PRA to make the task setting more concrete to the subject matter experts.

ATTACHMENT B-1

MATERIALS USED IN THE SLIM-MAUD TEST APPLICATION

1. Descriptions of Level A and B Tasks
2. Task Classification Procedures
3. SLIM-MAUD Facilitator's Guide
4. Introduction to SLIM-MAUD Session for First-time Users
5. Direct Estimate Response Instructions
6. Questionnaire for Evaluating SLIM-MAUD by Participants in SLIM-MAUD Test

ATTACHMENT B-1

1. DESCRIPTIONS OF LEVEL A AND B TASKS

LEVEL A TASKS

Assumptions to be used by judges in assessing Level A tasks:

- When reading the Level A tasks, everything that is not underlined is "given" and sets the stage for the underlined question.
- A senior reactor operator and a reactor operator are in the control room at all times.
- The persons performing the action in each task have been in their current job position for at least six months.
- No one involved in performing these tasks is wearing any type of protective clothing.
- The operator(s) does not have an unlimited amount of time in which to take action.

Level A Tasks:

(1) During a loss-of-off-site-power transient, several failures have rendered the high pressure coolant injection (HPCI) and the reactor core isolation cooling (RCIC) systems inoperable. Core cooling can be established with either low pressure coolant injection or low pressure core spray, but pressure must be reduced first. Procedural guidelines specify manual actuation of the automatic depressurization system (ADS) to reduce pressure. What is the likelihood that the operator will fail to actuate the ADS manually within 10 minutes?

(2) During a loss-of-off-site-power transient, the generator has tripped, the reactor has scrammed, and the normal feedwater system is inoprerable. According to the procedures, the reactor water level should be recovered and maintained by manually operating the reactor core isolation cooling (RCIC) system. What is the likelihood that the operator will fail to operate the RCIC system correctly?

(3) During a loss-of-off-site-power transient, the generator has tripped, the reactor has scrammed, and the normal feedwater system is inoperable. According to the emergency procedures, the operator must operate the nuclear instrumentation system by inserting the source and intermediate range monitors to verify that reactor power is decreasing following the scram. What is the likelihood that the operator will fail to operate the nuclear instrumentation system correctly?

(4) One of the main steam relief valves inadvertently opens. The operator, after successfully closing the valve, is monitoring the suppression pool temperature. The indicated temperature of the suppression pool is 95°F. According to procedures, this requires that the residual heat removal (RHR) system be manually placed in the suppression pool cooling mode. What is the likelihood that the operator will fail to actuate the suppression pool cooling mode of RHR?

(5) One of the main steam relief valves inadvertently opens. The operator mistakenly thinks he has reclosed the valve; however, the valve is still open. The operator properly places the RHR system in the suppression pool cooling mode when the temperature reaches 95°F. The temperature eventually reaches 110°F. The procedure then specifies that the operator must scram the reactor manually. What is the likelihood that the operator will fail to scram the reactor?

(6) A transient has occurred, the high pressure coolant injection (HPCI) system is operating, and the suppression pool cooling is inoperable. The operator notices that the HPCI system has inadvertently switched to suppression pool suction. The condensate storage tank (CST) level and the suppression pool level are both normal. The operator checks and finds that the CST water is still plentiful. What is the likelihood that the operator will not realize that high suppression pool temperature could ultimately fail HPCI due to loss of net positive suction head?

(7) A transient has occurred, the high pressure coolant injection (HPCI) system is operating, and the suppression pool cooling system is inoperable. The operator notices that the HPCI system has automatically switched to suppression pool suction. He checks and finds that the condensate storage tank (CST) water is still plentiful. The operator realizes that high suppression pool temperature could ultimately fail HPCI. What is the likelihood that he will fail to take the appropriate action to return the system manually so that the CST is the water supply?

(8) The plant is experiencing an extended station blackout (loss of on-site and off-site power) greater than 5 hours. Continued operation of the reactor core isolation cooling (RCIC) and high pressure coolant injection (HPCI) systems depends on sufficient room cooling for the equipment. What is the likelihood that the operator will fail to take precautions such as opening doors or providing other ventilation to ensure that the vital system equipment is being properly cooled?

(9) A transient has occurred, and the reactor has failed to scram. The operator, realizing what has happened, consults the emergency procedures for dealing with an anticipated transient without scram. The procedure states that he should attempt to trip the reactor manually. The operator attempts this but is unsuccessful. The procedure then calls for him to use the standby liquid control (SLC) system. What is the likelihood that the operator will fail to initiate SLC within 5-10 minutes after he reads the procedural step telling him to do so?

(10) A station blackout including total failure of the diesel generator system has just occurred. After the first immediate steps have been taken, the emergency procedures are referenced. What is the likelihood that the operator will attempt to restore off-site power before he attempts to restore power using the diesel generators?

(11) A transient has occurred, and the reactor protection system has failed to insert the rods. All attempts to manually scram the reactor have failed. According to the procedures, the operator is now required to manually insert the rods. What is the likelihood that the operator will fail to attempt to manually insert the rods using reactor manual control?

(12) A loss-of-coolant acident (LOCA) has occurred. The residual heat removal service water (RHRSW) system must be manually initiated within the first 30 minutes after the transient to obrtain successful long-term decay heat removal. The emergency operating procedures contain detailed instructions on operating the RHRSW. What is the likelihood that the operator will fail to recognize that he should initiate RHRSW within 30 minutes?

(13) A loss of coolant accident (LOCA) has occurred. The residual heat removal service water (RHRSW) system must be manually initiated to obtain successful long-term decay heat removal. The emergency operating procedures contain detailed instructions on operating the RHRSW, but the operator has so much to do he fails to operate the RHRSW. After 40 minutes, the operator gets a high suppression pool temperature alarm. What is the likelihood that he will then fail to diagnose the problem correctly and take steps to initiate RHRSW?

(14) The residual heat removal (RHR) system is providing shutdown cooling when the running RHR pump trips because of an electrical fault. The operator acknowledges that the pump tripped. Procedures state that the operator is to restore shutdown cooling. What is the likelihood that the operator will fail to attempt to restore RHR cooling within 10 minutes?

(15) The high pressure coolant injection (HPCI) system and the reactor core isolation cooling (RCIC) system have automatically initiated. The plant has experienced a total loss of instrument air. The pneumatic valves that control the cooling water to HPCI and RCIC room coolers do not open on demand because of the loss of instrument air. Opening these valves requires local operation. What is the likelihood that the operator will fail to open these valves within 1 hour?

LEVEL B TASKS

Assumptions to be used by judges in assessing Level B tasks:

- There is a one-man team in the control room during the performance of these tasks.
- These tasks take place during routine operations.
- The persons performing the action in each task have been in their current job position for at least six months.
- No one involved in performing these tasks is wearing any type of protective clothing.

Level B tasks:

(1) An operator chooses the wrong switch from a set of switches that all look similar and are identified only by labels.

(2) An operator chooses the wrong switch from a set of switches that all look similar and are grouped according to their functions.

(3) An operator chooses the wrong switch from a set of switches that all look similar and are arranged with clearly drawn mimic lines.

(4) The control room is designed so that they are moved to the right if the operator wants to turn on a component. The operator makes an error and turns a rotary control that has three or more positions to the left when he intends to turn the component on.

(5) Two or more locally operated valves are not clearly labeled. In addition, they are very similar in size and shape, they are in the same state (either open or closed), and they all have been tagged in a similar fashion. (The tags are all the same color, etc.) The operator attempts to place one of these valves back in service, but he mistakenly chooses the wrong one.

(6) A locally operated valve is clearly and unambiguously labeled and is not located near any similar-appearing valves. The operator intends to place the valve back in service, but he mistakenly chooses the wrong one.

(7) An operator reads the wrong meter in a group of meters that all look similar. They are arranged with clearly drawn mimic lines.

(8) An operator reads the wrong meter in a group of meters that all look similar. The meters are grouped according to their functions.

(9) An operator reads the wrong meter in a group of meters that all look similar and are identified only by labels.

(10) An equipment or auxiliary operator selects the wrong circuit breaker from a group of circuit breakers thast are located outside the control room. The circuit breakers are densely grouped and identified only by labels.

(11) A locally operated valve has a rising stem and a position indicator. An auxiliary operator, while using written procedures to check a valve lineup, fails to realize that the valve is not in its proper position after a maintenance person has performed a procedure intended to restore it to its proper position after maintenance.

(12) A meter has jammed so that the pointer is stuck on the scale. When an operator reads the meter, he fails to realize that it is jammed even though the value displayed is erroneous.

(13) An operator incorrectly reads information from a graph that is in a procedure.

(14) Assume that five annunciators are alarming. An operator fails to act on any of them.

(15) Assume that 10 annuciators have alarmed and an operator has responded to nine of them. The operator fails to act on the one remaining annunciator.

(16) An operator reads a digital indicator incorrectly.

(17) A chart recorder has normal bands indicated on the scale. An operator incorrectly interprets the value shown when he scans the recorder.

(18) A chart recorder does not have normal bands indicated on the scale. An operator incorrectly interprets the value shown when he scans the recorder.

(19) A meter has normal bands indicated on the scale. An operator does not notice that the meter is out of range after he performs an initial control room evaluation. No written materials are used.

(20) An operator intends to operate a 10-position rotary selector switch. He sets it to the wrong position.

ATTACHMENT B-1

2. TASK CLASSIFICATION PROCECURES

Introduction

The Engineering Analysis and Human Factors Group at BNL has been engaged in applying a methodology for generating human reliability data in the form of human error probabilities (HEPs) for nuclear power plant PRA. This effort has been divided into several stages, the first stage to be initiated with your participation. Your contribution to this effort is appreciated greatly.

Task Classification Exercise

You will be taking part in a classification exercise where certain operator actions in nuclear power plants will be evaluated. The exercise will be completed in two steps. In the first step you will be brought together with several other nuclear experts. The group of participating experts will then discuss a list of operator actions to ensure that everyone agrees on the meaning of these actions. In the second step you will again be brought together with other nuclear experts. Through open discussion the group will then create sub-groups of the actions from the complete list. Specific instructions for each of these steps together with the criterion for creating the sub-groups will be explained fully when the expert groups meet.

Presented below is a list of 30 tasks generic to all BWR plants. The list comprises <u>incorrect</u> actions that an operator might possibly take in a nuclear power plan. Each actions is part of the more complex behavioral sequences undertaken by an operator in performing his duties. The 30 tasks have been divided into two levels. Level A and Level B. Level A tasks combine power plant systems with human actions that present job duties. Level B tasks include those that combine equipment components with human actions and those involving controls and displays. For convenience when referring to a specific Level A task, each task has been given a brief name. Separate sets of assumptions are associated with each task level. Each set of assumptions appears immediately before the task descriptions, A then B, to which it applies.

To facilitate the group sessions you should familiarize yourself with the 30 tasks <u>before</u> the sessions are convened. Please do the following:

- Take a few minutes to familiarize yourself with all 30 tasks.
  - Be certain to review the assumptions associated with each task level.
  - Ample space has been provided between each task description so that you can freely record your thoughts and questions.

- The last page of this booklet contains a brief questionnaire asking for some information on your background. Complete the questionnaire before the group sessions.

- Bring this booklet to the group sessions.
  - Do not however, bring any reference materials to the group session.

Again, thank you for your contribution to this important task.

The experts received the following instructions:

Instructions

"In today's session you will be asked to discuss the operator actions contained in the list of 30 tasks you were already given to review. The purpose of this session is to arrive at shared agreement over the meaning and wording of these tasks. A view-graph of each task will be presented to the group in turn. As each task is presented you are to indicate whether the task is inadequately described or whether you do not understand the task. If a task is inadequately described or not fully understood by everyone, it will be discussed by the group. The discussion will continue until the group agrees on the meaning and proper wording of the task. Then, the next task will be presented and the same procedures will be followed. Do you have any questions?"

PAIRED COMPARISON RATINGS OF TASK SIMILARITY

Introduction

In the previous session agreement was reached on the meaning and wording of the list of the 30 human action tasks. This resulted in a list of revised tasks. In today's session you will be asked to discuss these revised tasks. The purpose of your discussion will be to judge the similarity between tasks. Your judgments will be used as a basis for grouping tasks. This grouping procedure is an essential first step in the methodology the BNL Engineering Analysis and Human Factors Group is applying to PRA work."

Task Similarity

You will be asked to judge all 30 tasks on their degree of similarity. Level B tasks will be judged separately from Level A.

Criterion of Similarity

A failure by an operator to complete any of the actions contained in the list of 30 tasks will be affected by certain key factors. Examples of typical factors that affect the likelihood of an operator failure include:

1. The time available to perform the task,
2. The complexity of the task, and
3. How well trained the operator is.

These factors, along with all other key factors affecting the task outcome, are known as performance shaping factors (PSFs). You will be asked to judge the similarity between tasks on the basis of the PSFs affecting task outcome. More specifically, those tasks for which the outcome is affected by the same PSFs in the same way will be judged to have a high degree of similarity.

Procedures

You will be judging the similarity of tasks taken in pairs. For each Level, the description of pairs of tasks will be presented on viewgraphs. You are to judge the degree of similarity between the tasks on the basis of whether or not the likelihood an operator will fail to complete the task is affected by the same PSFs acting together in the same way. You will rate the degree of similarity on a nine-point scale as follows:

1 - Indicates the tasks are not similar at all--the failure likelihood for each task is affected by entirely different PSFs.
9 - Indicates the tasks are very similar--the failure likelihood for each task is affected by the same PSFs acting in the same way.

You will be given a set of nine cards, numbered 1 to 9, to express your ratings. After you have reviewed a pair of tasks you should select the card whose number best expresses your judgment of their similarity. The card you select should not be shown or announced to the group, but placed face down until everyone has selected a card. When instructed to do so, hold up your card. Discrepancies between the individual ratings will then be discussed by the group until a consensus rating is arrived at. These procedures will then be repeated for all of the remaining pairs of tasks.

A few preliminary trials will be carried out to help you become familiar with these procedures.

Additional Guidelines

"Your judgments of task similarity should not be based upon superficial features of the tasks nor on the likelihood of the task outcome. Rather, your judgments should reflect your best assessment that tasks are similar because their likelihood of failure are affected by the same PSFs acting in the same way. In making your judgments you should assess the tasks globally, rather than performing an in-depth analysis of the details of each task. Finally, the tasks should be considered generic across BWR plants, not specific to a particular plant.

"Do you have any questions about what you are expected to do?"

ATTACHMENT B-1

3. SLIM-MAUD FACILITATOR'S GUIDE

1. INTRODUCTION

It is useful to have an individual who can act as a facilitator for SLIM-MAUD sessions, especially when the sessions are conducted with inexperienced user groups. For groups who are experienced with SLIM-MAUD applications, one of the group members can act as the facilitator.

1.1 Facilitator's Role

The facilitator is responsible for eliciting inputs from all group members, for attempting to resolve disagreements among group members, for guiding the group toward arriving at a shared definition of the tasks being assessed, and toward arriving at agreement over the PSFs presumed to affect the outcome of the tasks. Furthermore, it is especially important for the facilitator to encourage a cooperative group atmosphere, motivating group members to share information and to work jointly on the task assessment problem.

2. GETTING STARTED

The facilitator should try to avoid a "we must follow instructions" atmosphere at the outset and throughout the SLIM-MAUD session. The emphasis should be on guidance and motivation, not on rigid adherence to procedural steps. The group should be encouraged to get started with its assessment, rather than dwelling at length on minor procedural details. Typically, group members have no difficulty understanding the procedures once an assessment session actually begins. Any reluctance by the group about getting started with the assessment can usually be overcome if the facilitator stresses MAUD's built-in capabilities for helping the group to move through each stage of the assessment: by providing keyboard training on how to enter data; by providing menus of options clearly described in common language; by detecting errors and then actively assisting in editing or correcting them; and by helping the group to make other modifications when they wish to change their mind.

It is important that the facilitator engender an atmosphere of partnership between the facilitator, the group members, and MAUD for each assessment session. It should be made clear that MAUD and the facilitator are there to help the group, not to control it. Furthermore, group members should frequently be assured that they possess expertise which is valued and that the procedures are designed to draw out that expertise and bring it to bear on the assessments being made.

Before beginning to work directly with MAUD, it is useful if the facilitator compiles a preliminary list of PSFs and allows group members an opportunity to review the list. The list of PSFs, presumed to affect the outcome of the tasks to be assessed, can be used as a memory aid that can be referred to throughout the assessment session. It should be emphasized, however, that the group should not be bound to the list; each group must decide for itself whether to use any or all of the PSFs on the list, or whether to consider other PSFs not included in the list.

3. SELECTING AND RATING PSFs

MAUD works best when group members have concrete images in mind of the tasks they are assessing, and when they can clearly discriminate between those tasks (identifying similarities and differences) on the basis of the PSFs believed to affect the outcomes of the tasks. At the start of a SLIM-MAUD implementation triads of tasks are presented to the group. The facilitator should encourage members of the group to thoroughly discuss how the three tasks differ and to elaborate on those differences. This elaboration of task differences may take some time to complete, but this should not be of great concern because it helps group members to familiarize themselves with tasks and to get oriented to thinking about the tasks rather than about using MAUD. It is this elaboration of tasks that should lead to the selection of PSFs and to the specification of the poles on the PSF rating scale. Group members should be dissuaded from trying to choose PSFs first and then "fitting" them to the task under consideration.

Group members need not worry about excluding PSFs they think do not discriminate between the tasks or are irrelevant, nor those about which they have no expertise. The basic goal is to capture the expertise they possess. They should be the final arbiters of which PSFs to include for each subset of tasks assessed. Group members should be encouraged to develop their own idiosyncratic descriptions of generic PSFs so long as the descriptions are applicable to the tasks being assessed and the group is comfortable with the language of the descriptions.

After typing in descriptions of the tasks or of the PSFs selected, the group may decide a task or PSF description is unclear, too long for the display, or otherwise unsatisfactory. The group should be encouraged to change the task or PSF descriptions. To do so, it should answer "yes" the next time MAUD displays the prompt "Do you want to alter anything now? Answering "yes" shifts MAUD into its editing mode, providing the group with the opportunity to edit the offending task or PSF description, or to make whatever other changes may be desired.

Group members may initiate task-irrelevant conversations during MAUD sessions, especially when they find themselves in situations of stress, uncertainty, or conflict about task assessments. Interruptions can be costly because they divert the groups' attention from the task assessment situation and causes the members to lose the concrete images they have formed of the tasks. If permitted to continue, it will take some time to get the group to think about the tasks again.

When prolonged diversions occur, the facilitator should redirect the group back to the assessment objective, by reminding the members that they should be thinking about the tasks, by interjecting a comment or question about a specific task to be assessed, or by using some other means for changing the subject of conservation. It is usually easy to gain group support for this redirection because the majority of group members typically will prefer to get on with the assessment of tasks.

However, it is sometimes useful to permit comments about the system itself, or about other tangential matters in order to provide a break from the sometimes tedious job of assessing tasks. Such comments often occur at "natural breaks" in the procedures, such as when moving from PSF identification to PSF weighting, or when the assessment of a subject of tasks has been completed, or at the end of the session.

Sometimes a group will be dissatisfied with the names of the poles given to a PSF. In such instances, the facilitator should encourage the group to continue on, having it rate the actions being considered using these provisional pole names. Later on the poles can be renamed to the satisfaction of the group. If the group is still not satisfied with the provisional names, or with any other names that may be suggested, the entire rating dimension can be cancelled.

Users occasionally get into difficulty, especially in the early stages of the rating procedure, by giving names to the PSF poles that do not describe the opposites of a single bi-polar dimension. The group may need help here, and the facilitator can verbally hint at a name for the bi-polar opposite to the first name before the group enters the name of the second pole. Where the group has already entered pole names which refer to different dimensions (e.g., training required to stressful) the facilitator should point out that the group will run into difficulty when it attempts to rate the actions. The facilitator should explain (if the group doe not come to recognize it itself), that the chosen pole names refer to two dimensions, not one. As a result, each dimension requires its own separate set of ratings. Thus, two sets of ratings are required: a first set on a PSF having to do with training (ignoring stress) and a second set for a PSF having to do with stress (ignoring training). The facilitator should further point out that MAUD will give the group the opportunity to include both PSFs.

It is generally quite apparent when the group has selected all the PSFs it wishes to consider. This will occur at the point when none of the group members can think of (or persuade others to think of) another PSF about which any member has some expertise, or which can help to differentiate the tasks being assessed. Occasionally one or more group members will express apprehension about moving on to the next set of tasks to be assessed, thinking that an important PSF may have been left out. The facilitator should remind the group that it can always return any set of tasks to enter more PSFs if after the weighting procedure it is not satisfied with the results.

4. WEIGHTING THE PSFs

When performing the weighting procedure a group may encounter difficulty with the tradeoffs (i.e., moving an X all the way down the scale on a PSF until MAUD intervenes). This may arise because the group discovers a problem in interpreting the meaning of a PSF being displayed, or because it now thinks the ideal point is in the wrong place. The facilitator should review the meaning the group originally ascribed to the PSF, if that is the problem. The facilitator should also check to see if the displayed ideal point is correct, given the meaning of the PSF agreed to by the group. If the group still has difficulty making the required tradeoffs, the facilitator should remind it that MAUD provides the opportunity for editing the PSF or for modifying its ideal point. The group can pursue either of these options and then return to the weighting procedure.

5. AT THE END OF A SLIM-MAUD SESSION

After the SLIM-MAUD session is completed, the facilitator should review the results after the final summary is produced and ask the group if the rank-ordering of the SLI values produced are reasonable and acceptable. If the rank-orderings are acceptable, the procedure for converting the SLIs to probabilities can then be undertaken. If they are unacceptable, the group should be asked to identify the action which appears to be ordered incorrectly. The facilitator should then see if the group can identify a new PSF (i.e., one not yet entered) that discriminates between the offending action and the other actions considered, and that if added to the PSFs already considered could change the rank-ordering of the actions. The group can then enter the new PSF and examine the results to see if the revised rank-ordering is acceptable.

Before entering a new PSF, the group should be reminded that doing so will require it to go through the weighting procedure all over again in order to get a revised SLI ordering. If the group agrees to this, it should proceed with entering the new PSF and with the reweighting procedure which, when completed,will produce a revised summary of results. In principle a group can continue to make revisions as many times as it deems necessary to arrive at an acceptable SLI ordering. In practice, however, groups seldom make repeated revisions.

ATTACHMENT B-1

4. INTRODUCTION TO SLIM-MAUD SESSION FOR FIRST-TIME USERS

Estimating HEPs for input into HRA segments of PRAs has been a major goal of the NRC. Because of the extreme difficulty of collecting actual data on human errors, attempts have been made to develop methodologies that estimate HEPs using expert judges.

Today we will be using one of these methodologies. The methodology is SLIM-MAUD, which is a double acronym standing for Success Likelihood Index Methodology, on the one hand, and Multi-Attribute Utility Decomposition, on the other. SLIM, the underlying methodology, is a basic technique for eliciting your expertise and organizing it in a way that it can be used to make HEP estimates. The emphasis of SLIM is to bring your wide ranging knowledge and expertise to bear on a problem. MAUD is a software-based decision aid for implementing SLIM.

Performance Shaping Factors

The key underlying assumption of SLIM is that the success or failure of any human task in a particular situation depends upon the combined effects of a relatively small set of PSFs. In brief, PSFs include both human traits and conditions of the work setting that may influence an individual's performance.

Listed on the board are some representative PSFs that could influence the outcome of human tasks or actions in a given situation. Briefly review the PSFs. We will return to them in a moment. Today, with the help of SLIM-MAUD, we will identify the important PSFs that influence the outcome of human tasks or actions on the list you were given to review.

Procedures

This will be accomplished with a three-step procedure:

1. The SLIM-MAUD program will first elicit what you think are the most important PSFs influencing the outcome of the tasks (actions).

2. The the PSFs will be rated in importance.

3. Finally, the PSFs will be weighted with respect to one another.

Elicitation

With the SLIM-MAUD program the elicitation and rating of PSFs (Steps 1 and 2) actually takes place in a single step.

SLIM-MAUD program will ask you to enter not less than 4 and not more than 10 tasks to be evaluated. Here, tasks will be entered according to the subsets they are organized into (three subsets of Level A tasks and three subsets of Level B).

Then the MAUD program will present the tasks, three at a time, and elicit the PSFs that you think influence the outcome of the task. You should openly discuss the task until you can agree. It then asks you to rate, on a scale of 1 to 9, each of the tasks with respect to the identified PSFs.

Then MAUD will ask you to give the "ideal point" on each of the PSF scales. (This is done because some of the scales may not be linear, in which case the ideal point will not be equal to the scale's endpoint.)

The scale is then converted to a standard scale whose range is from 0 to 1 (based on the ideal point you previously gave it for the scales).

Weighting

After the rating procedure is completed, MAUD will have you weight each of the PSFs with respect to its relative importance in contributing to success or failure of the task. The weighting will be accomplished by having you make a series of comparisons between best and worst points on the scales you created. MAUD then computes the weightings on a scale that is normalized to sum to 1.

SLI

The ultimate goal of this whole process is to produce an index which can then be converted to a probability, an HEP. We call this index the success likelihood index or SLI. It is obtained by multiplying your ratings times your weightings on each course of action and summing the results. The index ranges in value from 0 to 1.

Zero means that all the PSFs on the action being considered are the worst or tied for the worst compared to the other.	One means that all the PSFs on the action being considered are the best or tied for best compared to the other.

Summary

Thus, here is the procedure in a nutshell.

1. You will input tasks into the SLIM-MAUD program in subsets.

2. Using MAUD, you will (1) identify PSFs, (2) rate them in importance, and (3) choose the scale's ideal point.

3. Then, using MAUD you will weight the PSFs.

4. The MAUD program will then take your inputs and compute the SLIs in each subset.

5. Finally, we will review the calculated SLIs to see if they are reasonable. If people disagree with the relative SLI, we will discuss and reassess tasks. When the SLIs are acceptable to the group, we will go on to the next subset of tasks.

Some Additional Guidance

1. Although the procedures may sound a little complicated, they are not, but feel free to ask questions. MAUD is very easy to use and will assist you in your efforts to assess the importance of PSFs on task outcomes.

2. The whole procedure is based upon its ability to thoroughly tap your expertise--to draw out your knowledge of factors that influence human performance.

   - Accordingly, the emphasis is on obtaining input from all members of the group. Obtaining this is best accomplished through co-operation where you exchange information freely among yourselves.

   - Think of your own knowledge as an important resource to be shared by the group.

   - Try to use the information as effectively as you can in making your assessments.

ATTACHMENT B-1

5. DIRECT ESTIMATE RESPONSE INSTRUCTIONS

NAME: ____

INSTRUCTIONS FOR COMPLETION OF DIRECT ESTIMATE AND UNCERTAINTY BOUNDARY JUDGMENTS

We would now like you to make direct probability estimates of the same set of Level A and Level B tasks which you evaluated with SLIM-MAUD. In addition, we would like your estimates of the upper and lower bounds of your probability estimates.

A single task is presented on each of the following sheets. Once you have read and understood each task on the left side of the page, put an X on the point on the scale on the right that represents your best estimate of the chance of the incorrect action occurring. Remember, you are to assume that the operator does not have an unlimited amount of time in which to take action.

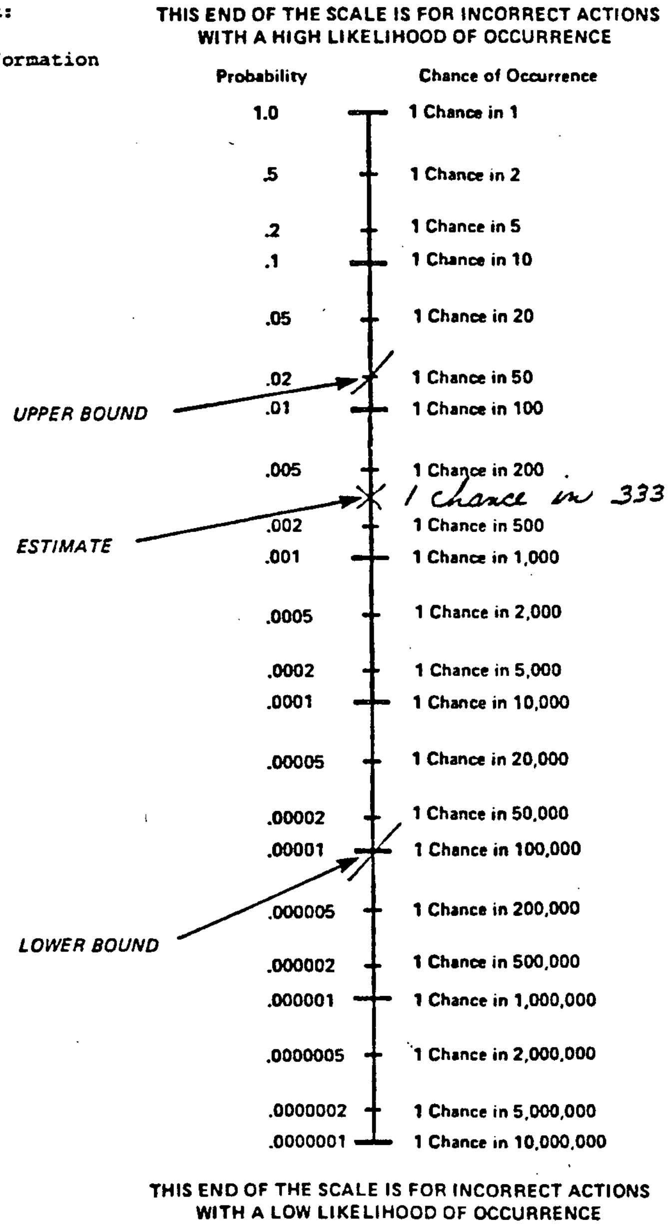
Next, place slash marks to indicate upper and lower bounds so that you are 90 percent certain that the value will fall within those bounds. If a mark or exact value that represents your estimate does not appear on the scale (e.g., 1 chance in 3,500), place your X or slash at the approximate position on the scale and write your estimate to the right of the scale.

A completed example follows these instructions which you can use to assist you in making your estimates.

Example of Completed Direct Estimate*

Estimate the chances that:

An operator will read information from a graph incorrectly.



*Adopted from Comer et al. (1984), Volume 2, Appendix B.

ATTACHMENT B-1

6. QUESTIONNAIRE FOR EVALUATING SLIM-MAUD BY PARTICIPANTS IN SLIM-MAUD TEST BACKGROUND INFORMATION

Please provide the information requested below. Be as accurate and complete as possible.

1. Your name: ______________________________

2. Highest level of education attained (circle one):

   a. HIGH SCHOOL DIPLOMA OR EQUIVALENT
   b. TRADE SCHOOL OR ASSOCIATE DEGREE
   c. BACHELOR'S DEGREE
   d. MASTER'S DEGREE
   e. OTHER: ______________________________ (Describe)

   Major Field: ______________________________
   Major Field: ______________________________
   Major Field: ______________________________

3. What is your nuclear power plant experience? (Indicate number of years in the blanks provided:)

   a. MILITARY: __________ years
   b. FOSSIL (COMMERCIAL): __________ years
   c. NUCLEAR (COMMERCIAL): __________ years
   d. OTHER: __________ years: ______________________________ (Describe the type of experience.)

4. With which type of plant are you experienced? (Circle one:)

   a. BWR
   b. PWR
   c. BOTH: ______________________________ (Describe)

5. Have you had simulator training? (Circle one:)

   a. YES - BWR
   b. YES - PWR

6. Do you hold any kind of license or certification? (Circle one:)

   a. YES
   b. NO

   (If yes), describe: ______________________________

7. Is there any other information we may have overlooked that would be useful in gaining an understanding of your nuclear power plant experience? Use the space below to provide such information.

____
____
____
____

The following questions are concerned with your experience in using SLIM-MAUD to estimate HEPs. Please answer all questions as accurately as you can.

(Please circle your answers)

1. The SLIM-MAUD procedures were easy to understand.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
2. The SLIM-MAUD procedures could be clearly followed.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
3. The SLIM-MAUD procedures effectively used my knowledge and experience.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
4. The SLIM-MAUD procedures seemed orderly.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
5. Using SLIM-MAUD improved my understanding of the factors that influence the likelihood of an incorrect action.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
6. The SLIM-MAUD procedures led to results that seemed meaningful.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
7. It would be easy to determine which PSFs had the greatest impact on HEP estimates by reviewing the SLIM-MAUD log of results.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
8. It was easy to estimate uncertainty bounds.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
9. My judgments of uncertainty bounds are accurate.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE
10. SLIM-MAUD can be useful to HRA segments of PRAs.	STRONGLY AGREE	AGREE	NEUTRAL	DIS-AGREE	STRONGLY DISAGREE

11. What is the likelihood that SLIM-MAUD will be acceptable for use by government and research PRA experts? (Circle number of your answer.)

    1. VERY LIKELY
    2. SOMEWHAT LIKELY
    3. NEITHER LIKELY OR UNLIKELY
    4. SOMEWHAT UNLIKELY
    5. VERY UNLIKELY

12. What is the likelihood that SLIM-MAUD will be acceptable for use by utility PRA experts? (Circle number of your answer.)

    1. VERY LIKELY
    2. SOMEWHAT LIKELY
    3. NEITHER LIKELY OR UNLIKELY
    4. SOMEWHAT UNLIKELY
    5. VERY UNLIKELY

Any additional comments you may have regarding your experience in using SLIM-MAUD will be appreciated. Write your comments in the space provided below.

ATTACHMENT B-2

DOCUMENTATION OF SLIM-MAUD SESSIONS

1. Summary of Reports of SLIM-MAUD Sessions

2. Technical Recorder's Notes of Contents of Group Discussions During SLIM-MAUD Sessions

ATTACHMENT B-2

1. SUMMARY OF REPORTS OF SLIM-MAUD SESSIONS

- GROUP: 1
- TASKS: LEVEL A, SUBSET 1

SUMMARY OF SESSION GROUP1 LEVEL A TASKS SUBSET SO FAR:

Current order of assessment of Likelihood of Success of actions
from best to worst
(Success Likelihood Indices are given in brackets)

SP COOLING AFTER MSRV OPENS (0.80) (BEST

RHR SD COOL RESTORE (0.75)

NI OPERATION AFTER LOSP (0.44)

HI SP T-PREV HPCI FAIL (0.30)

MAN SCRAM ON HI SP T (0.29)

RCIC OPERATION AFTER LOSP (0.24) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	RCIC OPERATION AFTER LOSP	NI OPERATION AFTER LOSP	SP COOLING AFTER MSRV OPENS	MAN SCRAM ON HI SP T	HI SP T-PREV HPCI FAIL	RHR SD COOL RESTORE	Performance Shaping Factor
(1)	7	7	3	6	5	2	LOW STRESS(1) to HI STRESS(9) Ideal value= 3
(2)	4	2	3	1	1	2	LONG TIME(1) to SHORT TIME(9) Ideal value= 2
(5)	8	2	6	6	7	4	LESS TRAINING(1) to MORE TRAINING (9) Ideal value= 1
(6)	3	1	3	6	7	4	LOW CONSEQUENCES(1) to HI CONSEQUENCES(9) Ideal value= 3

* The following information shows your assessment of Likelihood of Success for the actions under consideration

1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	RCIC OPERATION AFTER LOSP	NI OPERATION AFTER LOSP	SP COOLING AFTER MSRV OPENS	MAN SCRAM ON HI SP T	HI SP T-PREV HPCI FAIL	RHR SD COOL RESTORE	Performance Shaping Factor
(1)	0.00	0.00	1.00	0.25	0.50	0.75	LOW STRESS to HI STRESS relative importance = 0.43
(2)	0.00	1.00	0.50	0.50	0.50	1.00	LONG TIME to SHORT TIME relative importance = 0.08
(5)	0.00	1.00	0.33	0.33	0.17	0.67	LESS TRAINING to MORE TRAINING relative importance = 0.24
(6)	1.00	0.50	1.00	0.25	0.00	0.75	LOW CONSEQUENCES to HI CONSEQUENCES relative importance = 0.24

The following scales are no longer in use for the reasons given below

(3) LESS TRAINING to MORE TRAINING
This scale was cancelled because of its
similarity with scale number 4

(4) LOW COMPLEXITY to HIGH COMPLEXITY
This scale was cancelled because of its
similarity with scale number 3

END OF SUMMARY.

GROUP: 1
TASKS: LEVEL A, SUBSET 2

SUMMARY OF SESSION GROUP1 LEVEL A TASKS SUBSET 2 SO FAR:

Current order of assessment of Likelihood of Success of actions
from best to worst
(Success Likelihood Indices are given in brackets)

LOCA W/O SP COOL (0.83) (BEST

LOCA-INIT RHRSW (0.57)

HI SP T-POT HPCI FAIL (0.36)

STA BLA-RM COOL/H-R (0.24)

IA LOSS-LOC OPS VLV (0.15) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	HI SP T-POT HPCI FAIL	STA BLA-RM COOL/H-R	LOCA-INIT RHRSW	LOCA W/O SP COOL	IA LOSS-LOC OPS VLV	Performance Shaping Factor
(2)	3	1	6	9	2	LOW CUE(1) to HI CUE(9) Ideal value= 9
(3)	5	7	8	7	7	LOW STRESS(1) to HI STRESS(9) Ideal value= 3
(4)	7	6	3	2	8	LOW TRAINING(1) to HI TRAINING(9) Ideal value= 1
(5)	9	7	3	2	8	POOR PROC(1) to GOOD PROC(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

SUMMARY OF SESSION GROUP1 LEVEL A TASKS SUBSET 2 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

LOCA W/O SP COOL (0.83) (BEST

LOCA-INIT RHRSW (0.57)

HI SP T-POT HPCI FAIL (0.36)

STA BLA-RM COOL/H-R (0.24)

IA LOSS-LOC OPS VLV (0.15) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	HI SP T-POT HPCI FAIL	STA BLA-RM COOL/H-R	LOCA-INIT RHRSW	LOCA W/O SP COOL	IA LOSS-LOC OPS VLV	Performance Shaping Factor
(2)	0.25	0.00	0.63	1.00	0.13	LOW CUE to HI CUE relative importance = 0.24
(3)	1.00	0.33	0.00	0.33	0.33	LOW STRESS to HI STRESS relative importance = 0.26
(4)	0.17	0.33	0.83	1.00	0.00	LOW TRAINING to HI TRAINING relative importance = 0.25
(5)	0.00	0.29	0.86	1.00	0.14	POOR PROC to GOOD PROC relative importance = 0.25

The following scales are no longer in use for the reasons given below

(1) LOW INFO to HI INFO
This scale was cancelled because there was very little difference between your preferences for actions on it
END OF SUMMARY.

GROUP: 1
TASKS: LEVEL A, SUBSET 3

SUMMARY OF SESSION GROUP1 LEVEL A TASKS SUBSET 3 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

ATWS-MAN INS RODS (0.82) (BEST

MAN ADS ACT/LOSP (0.77)

ATWS-INIT SLC (0.40)

STA BLA-OS POW REST (0.28) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	MAN ADS ACT/LOSP	ATWS-INIT SLC	STA BLA-OS POW REST	ATWS-MAN INS RODS	Performance Shaping Factor
(1)	7	8	7	6	LOW STRESS(1) to HI STRESS(9) Ideal value= 3
(2)	7	5	3	7	SHORT TIME AVAIL(1) to LONG TIME AVAIL(9) Ideal value= 7
(3)	3	5	2	1	LOW TRAINING NEC(1) to HIGH TRAINING NEC(9) Ideal value= 1
(4)	2	2	5	4	LOW COMPLEX(1) to HI COMPLEX(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	MAN ADS ACT/LOSP	ATWS-INIT SLC	STA BLA-OS POW REST	ATWS-MAN INS RODS	Performance Shaping Factor
(1)	0.50	0.00	0.50	1.00	LOW STRESS to HI STRESS relative importance = 0.27
(2)	1.00	0.50	0.00	1.00	SHORT TIME AVAIL to LONG TIME AVAIL relative importance = 0.27
(3)	0.50	0.00	0.75	1.00	LOW TRAINING NEC to HIGH TRAINING NEC relative importance = 0.20
(4)	1.00	1.00	0.00	0.33	LOW COMPLEX to HI COMPLEX relative importance = 0.27

END OF SUMMARY.

GROUP: I
TASKS: LEVEL B ,SUBSET I

SUMMARY OF SESSION GROUP1 LEVEL B TASKS SUBSET 1 SO FAR:

Current order of assessment of Likelihood of Success of actions
from best to worst
(Success Likelihood Indices are given in brackets)

METER W/NB-OUT RGN-READ AS IS (1.00) <BEST

WRONG METER W/MIMICS (0.90)

WRONG SWITCH W/MIMICS (0.82)

WRONG METER BY FUNCTION (0.46)

JAMMED METER/READ AS IS (0.00) <WORST

Ratings of actions on the scales you are currently using

Rating scale number	WRONG SWITCH W/MIMICS	WRONG METER W/MIMICS	WRONG METER BY FUNCTION	JAMMED METER/READ AS IS	METER W/NB-OUT RGN-READ AS IS	Performance Shaping Factor
(1)	3	3	5	8	2	LOW COMPLEX(1) to HI COMPLEX(9) Ideal value= 1
(2)	4	3	6	8	3	LOW TRAINING NEC(1) to HI TRAINING NEC(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	WRONG SWITCH W/MIMICS	WRONG METER W/MIMICS	WRONG METER BY FUNCTION	JAMMED METER/READ AS IS	METER W/NB-OUT RGN-READ AS IS	Performance Shaping Factor
(1)	0.83	0.83	0.50	0.00	1.00	LOW COMPLEX to HI COMPLEX relative importance = 0.57
(2)	0.80	1.00	0.40	0.00	1.00	LOW TRAINING NEC to HI TRAINING NEC relative importance = 0.43

END OF SUMMARY.

GROUP: 1
TASKS: LEVEL B , SUBSET 2

SUMMARY OF SESSION GROUP1 LEVEL B TASKS SUBSET 2 SO FAR:

Current order of assessment of Likelihood of Success of actions
from best to worst
(Success Likelihood Indices are given in brackets)

CHT REC W/NB-WRONG READ (1.00) (BEST

WRONG METER W/LABELS (0.41)

WRONG SWITCH BY FUNCTION (0.33)

WRONG CB W/LABELS (0.31)

CHT REC W/O NB WRONG READ (0.28)

WRONG SWITCH W/LABELS (0.16) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	WRONG SWITCH W/LABELS	WRONG SWITCH BY FUNCTION	WRONG METER W/LABELS	WRONG CB W/LABELS	CHT REC W/NB-WRONG READ	CHT REC W/O NB WRONG READ	Performance Shaping Factor
(1)	5	4	4	5	1	4	LOW COMPLEXITY(1) to HIGH COMPLEXITY(9) Ideal value= 1
(2)	6	5	5	6	3	4	LOW EXPERIENCE(1) to HI EXPERIENCE (9) Ideal value= 1
(3)	5	5	4	3	2	7	LOW V A(1) to HIGH V A(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	WRONG SWITCH W/LABELS	WRONG SWITCH BY FUNCTION	WRONG METER W/LABELS	WRONG CB W/LABELS	CHT REC W/NB-WRONG READ	CHT REC W/O NB WRONG READ	Performance Shaping Factor
(1)	0.00	0.25	0.25	0.00	1.00	0.25	LOW COMPLEXITY to HIGH COMPLEXITY relative importance = 0.30
(2)	0.00	0.33	0.33	0.00	1.00	0.67	LOW EXPERIENCE to HI EXPERIENCE relative importance = 0.31
(3)	0.40	0.40	0.60	0.80	1.00	0.00	LOW V A to HIGH V A relative importance = 0.39

END OF SUMMARY.

GROUP: 1
TASKS: LEVEL B, SUBSET 3

SUMMARY OF SESSION GROUP1 LEVEL B TASKS SUBSET 3 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

5 ALARM/NO ACT (1.00) (BEST

WRONG VLV/DUMMY (0.92)

10 ALARM/9 ACT (0.39)

WRONG VLV/COIN FLIP (0.00) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	WRONG VLV/COIN FLIP	WRONG VLV/DUMMY	5 ALARM/NO ACT	10 ALARM/9 ACT	Performance Shaping Factor
(1)	9	2	1	7	LOW M A(1) to HI M A(9) Ideal value= 1
(2)	8	3	3	5	LOW ENV IMPACT(1) to HI ENV IMPACT(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	WRONG VLV/COIN FLIP	WRONG VLV/DUMMY	5 ALARM/NO ACT	10 ALARM/9 ACT	Performance Shaping Factor
(1)	0.00	0.88	1.00	0.25	LOW M A to HI M A relative importance = 0.60
(2)	0.00	1.00	1.00	0.60	LOW ENV IMPACT to HI ENV IMPACT relative importance = 0.40

END OF SUMMARY.

GROUP: 2
TASKS: LEVEL A, SUBSET 1

SUMMARY OF SESSION LEVEL A SUBSET 1 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

NI OP AFTER LOSP (0.61) (BEST

RHR SD COOLING RES (0.59)

HI SP TEMP STOP HPCI FAIL (0.57)

RCIC OP AFTER LOSP (0.53)

MAN SCRAM ON HI SP TEMP (0.50)

SP COOL AFTER MRSV OPEN (0.46) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	RCIC OP AFTER LOSP	NI OP AFTER LOSP	SP COOL AFTER MRSV OPEN	MAN SCRAM ON HI SP TEMP	HI SP TEMP STOP HPCI FAIL	RHR SD COOLING RES	Performance Shaping Factor
(1)	8	9	6	8	4	7	REQUIRES DIAGNOSIS(1) to MANDATED BY PROCEDURES(9) Ideal value= 9
(2)	5	7	5	2	7	9	HIGH STRESS LEVEL(1) to LOW STRESS LEVEL(9) Ideal value= 9
(3)	3	2	4	1	8	8	HIGH TIME PRESSURE(1) to LOW TIME PRESSURE(9) Ideal value= 6
(4)	4	6	5	4	3	5	SEVERE CONSEQUENCES(1) to MINOR CONSEQUENCES(9) Ideal value= 2
(5)	3	7	2	9	5	3	COMPLEX(1) to SIMPLE(9) Ideal value= 9
(6)	8	9	7	9	7	5	REQUIRES TEAMWORK(1) to NO TEAMWORK REQUIRED(9) Ideal value= 7

* The following information shows your assessment of Likelihood of Success for the actions under consideration

1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	RCIC OP AFTER LOSP	NI OP AFTER LOSP	SP COOL AFTER MRSV OPEN	MAN SCRAM ON HI SP TEMP	HI SP TEMP STOP HPCI FAIL	RHR SD COOLING RES	Performance Shaping Factor
(1)	0.80	1.00	0.40	0.80	0.00	0.60	REQUIRES DIAGNOSIS to MANDATED BY PROCEDURES relative importance = 0.25
(2)	0.43	0.71	0.43	0.00	0.71	1.00	HIGH STRESS LEVEL to LOW STRESS LEVEL relative importance = 0.16
(3)	0.67	0.33	1.00	0.00	1.00	1.00	HIGH TIME PRESSURE to LOW TIME PRESSURE relative importance = 0.21
(4)	0.67	0.00	0.33	0.67	1.00	0.33	SEVERE CONSEQUENCES to MINOR CONSEQUENCES relative importance = 0.10
(5)	0.14	0.71	0.00	1.00	0.43	0.14	COMPLEX to SIMPLE relative importance = 0.24
(6)	0.50	0.00	1.00	0.00	1.00	0.00	REQUIRES TEAMWORK to NO TEAMWORK REQUIRED relative importance = 0.04

END OF SUMMARY.

SUMMARY OF SESSION LEVEL A SUBSET 2 SO FAR:

GROUP: 2
TASKS: LEVEL A ,SUBSET 2

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

LOCA INIT RHRSW (0.70) (BEST

LOCA W NO SP COOL INIT RHRSW (0.68)

HI SP TEMP REALIZE HPCI FAIL (0.56)

LOSS IA LOCAL V OP (0.37)

STA BO RM COOL TO HPCI RCIC (0.07) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	HI SP TEMP REALIZE HPCI FAIL	STA BO RM COOL TO HPCI RCIC	LOCA INIT RHRSW	LOCA W NO SP COOL INIT RHRSW	LOSS IA LOCAL V OP	Performance Shaping Factor
(4)	8	7	3	2	7	MANDATED BY PROCEDURES(1) to REQUIRE DIAGNOSIS(9) Ideal value= 1
(5)	2	8	3	3	8	INDIVIDUAL(1) to COORDINATED(9) Ideal value= 1
(7)	5	1	3	2	4	STRESSFUL CONSEQUENCES(1) to MINOR CONSEQUENCES(9) Ideal value= 7

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	HI SP TEMP REALIZE HPCI FAIL	STA BORM COOL TO HPCI RCIC	LOCA INIT RHRSW	LOCA W NO SP COOL INIT RHRSW	LOSS IA LOCAL V OP	Performance Shaping Factor
(4)	0.00	0.17	0.83	1.00	0.17	MANDATED BY PROCEDURES to REQUIRE DIAGNOSIS relative importance = 0.44
(5)	1.00	0.00	0.83	0.83	0.00	INDIVIDUAL to COORDINATED relative importance = 0.17
(7)	1.00	0.00	0.50	0.25	0.75	STRESSFUL CONSEQUENCES to MINOR CONSEQUENCES relative importance = 0.39

The following scales are no longer in use for the reasons given below

(1) MINIMUM TRAINING REQ to MAXIMUM TRAINING REQ
This scale was cancelled because of its
similarity with scale number 3
(2) SEVERE CONSEQUENCES to MINOR CONSEQUENCES
This scale was cancelled because of its
similarity with scale number 6
(3) REQUIRE DIAGNOSIS to MANDATED BY PROCEDURES
This scale was cancelled because of its
similarity with scale number 1
(6) HIGH STRESS to LOW STRESS
This scale was cancelled because of its
similarity with scale number 2
END OF SUMMARY.

SUMMARY OF SESSION LEVEL A SUBSET 3 SO FAR:

GROUP: 2
TASKS: LEVEL A, SUBSET 3

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

MAN ADS ACT AFTER LOSP (0.91) (BEST

ATWS INIT SLC (0.63)

ATWS MAN INSERT RODS (0.44)

STA BO OSP RESTORE (0.15) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	MAN ADS ACT AFTER LOSP	ATWS INIT SLC	STA BO OSP RESTORE	ATWS MAN INSERT RODS	Performance Shaping Factor
(1)	2	2	8	2	INDIVIDUAL(1) to COORDINATED(9) Ideal value= 1
(2)	6	1	3	2	HIGH STRESS(1) to MINIMAL STRESS(9) Ideal value= 8
(3)	7	8	3	5	COMPLEX(1) to SIMPLE(9) Ideal value= 9

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	MAN ADS ACT AFTER LOSP	ATWS INIT SLC	STA BO OSP RESTORE	ATWS MAN INSERT RODS	Performance Shaping Factor
(1)	1.00	1.00	0.00	1.00	INDIVIDUAL to COORDINATED relative importance = 0.20
(2)	1.00	0.00	0.40	0.20	HIGH STRESS to MINIMAL STRESS relative importance = 0.38
(3)	0.80	1.00	0.00	0.40	COMPLEX to SIMPLE relative importance = 0.43

END OF SUMMARY.

GROUP: 2
TASKS: LEVEL B , SUBSET 1

SUMMARY OF SESSION LEVEL B SUBSET 1 SO FAR:

Current order of assessment of Likelihood of Success of actions
from best to worst
(Success Likelihood Indices are given in brackets)

WRONG SW W MIMICS (1.00) (BEST

WRONG METER W MIMICS (0.85)

OUT OF RANGE METER (0.58)

WRONG METER BY FUNCTION (0.46)

STUCK METER (0.00) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	WRONG SW W MIMICS	WRONG METER W MIMICS	WRONG METER BY FUNCTION	STUCK METER	OUT OF RANGE METER	Performance Shaping Factor
(2)	3	3	5	7	3	GOOD HUMAN FACTORS(1) to POOR HUMAN FACTORS(9) Ideal value= 1
(3)	8	8	5	1	3	HIGH TRAINING REQ(1) to LESS TRAINING REQ (9) Ideal value= 9
(4)	1	8	8	9	9	MANUAL(1) to COGNITIVE(9) Ideal value= 2

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	WRONG SW W MIMICS	WRONG METER W MIMICS	WRONG METER BY FUNCTION	STUCK METER	OUT OF RANGE METER	Performance Shaping Factor
(2)	1.00	1.00	0.50	0.00	1.00	GOOD HUMAN FACTORS to POOR HUMAN FACTORS relative importance = 0.48
(3)	1.00	1.00	0.57	0.00	0.29	HIGH TRAINING REQ to LESS TRAINING REQ relative importance = 0.34
(4)	1.00	0.17	0.17	0.00	0.00	MANUAL to COGNITIVE relative importance = 0.18

The following scales are no longer in use for the reasons given below

(1) MIMICS to NO MIMICS
You cancelled this scale after trying to rate the actions on it.
END OF SUMMARY.

GROUP: 2
TASKS: LEVEL B, SUBSET 2

SUMMARY OF SESSION LEVEL B SUBSET 2 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

WRONG SW W FUNCTION (0.90) (BEST

WRONG SW W LABELS (0.83)

WRONG METER W LABELS (0.65)

CH RECORDER W NORM BANDS (0.57)

CH RECORDER WO BANDS (0.29)

AOP WRONG CB W LABELS (0.21) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	WRONG SW W LABELS	WRONG SW W FUNCTION	WRONG METER W LABELS	AOP WRONG CB W LABELS	CH RECORDER W NORM BANDS	CH RECORDER WO BANDS	Performance Shaping Factor
(1)	7	4	7	9	3	7	GOOD HUMAN FACTORS(1) to POOR HUMAN FACTORS(9) Ideal value= 1
(2)	1	1	1	6	1	1	INDIVIDUAL(1) to COORDINATED(9) Ideal value= 1
(3)	1	1	8	6	9	9	MANUAL(1) to COGNITIVE(9) Ideal value= 2
(4)	8	7	8	4	4	2	HIGH TRAINING REQ(1) to LESS TRAINING REQ(9) Ideal value= 9

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	WRONG SW W LABELS	WRONG SW W FUNCTION	WRONG METER W LABELS	AOP WRONG CB W LABELS	CH RECORDER W NORM BANDS	CH RECORDER WO BANDS	Performance Shaping Factor
(1)	0.33	0.83	0.33	0.00	1.00	0.33	GOOD HUMAN FACTORS to POOR HUMAN FACTORS relative importance = 0.26
(2)	1.00	1.00	1.00	0.00	1.00	1.00	INDIVIDUAL to COORDINATED relative importance = 0.20
(3)	1.00	1.00	0.17	0.50	0.00	0.00	MANUAL to COGNITIVE relative importance = 0.22
(4)	1.00	0.83	1.00	0.33	0.33	0.00	HIGH TRAINING REQ to LESS TRAINING REQ relative importance = 0.32

END OF SUMMARY.

GROUP: 2
TASKS: LEVEL B, SUBSET 3

SUMMARY OF SESSION LEVEL B SUBSET 3 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

FIVE ANNUNCIATORS ON (0.73) (BEST

SINGLE VALVE (0.49)

10 ANNUNCIATORS ON (0.45)

ONE OF MANY VALVES (0.08) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	ONE OF MANY VALVES	SINGLE VALVE	FIVE ANNUNCIATORS ON	10 ANNUNCIATORS ON	Performance Shaping Factor
(1)	9	1	8	7	GOOD HUMAN FACTORS(1) to POOR HUMAN FACTORS(9) Ideal value= 1
(2)	4	7	8	8	HIGH TRAINING REQ(1) to LESS TRAINING REQ(9) Ideal value= 9
(3)	6	6	1	1	INDIVIDUAL(1) to COORDINATED(9) Ideal value= 1
(4)	3	3	6	2	HIGH WORKLOAD(1) to LOW WORKLOAD(9) Ideal value= 9

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	ONE OF MANY VALVES	SINGLE VALVE	FIVE ANNUNCIATORS ON	10 ANNUNCIATORS ON	Performance Shaping Factor
(1)	0.00	1.00	0.13	0.25	GOOD HUMAN FACTORS to POOR HUMAN FACTORS relative importance = 0.31
(2)	0.00	0.75	1.00	1.00	HIGH TRAINING REQ to LESS TRAINING REQ relative importance = 0.13
(3)	0.00	0.00	1.00	1.00	INDIVIDUAL to COORDINATED relative importance = 0.24
(4)	0.25	0.25	1.00	0.00	HIGH WORKLOAD to LOW WORKLOAD relative importance = 0.32

END OF SUMMARY.

SUMMARY OF SESSION G3LAS1 SO FAR:

GROUP: 3
TASKS: LEVEL A, SUBSET 1

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

RHR S/D COOL/RESTORE (0.83) (BEST

SP COOL AFTER MSRV OPEN (0.67)

MAN RX SCRAM ON HI SP TEMP (0.67)

RCIC OP AFTER LOSP (0.63)

NI OP AFTER LOSP (0.44)

HI SP TEMP/PREV HPCI FAIL (0.33) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	RCIC OP AFTER LOSP	NI OP AFTER LOSP	SP COOL AFTER MSRV OPEN	MAN RX SCRAM ON HI SP TEMP	HI SP TEMP/PREV HPCI FAIL	RHR S/D COOL/RESTORE	Performance Shaping Factor
(1)	7	9	5	1	6	5	LITTLE TIME AVAILABLE(1) to MORE TIME AVAILABLE(9) Ideal value= 6
(2)	8	9	7	1	7	6	HIGHER CONSEQUENCE(1) to LOWER CONSEQUENCE(9) Ideal value= 3
(4)	8	8	6	7	1	6	MORE DIAGNOSIS REQ(1) to LESS DIAGNOSIS REQ(9) Ideal value= 5
(5)	2	2	5	1	8	3	FREQ SIM PRACTICE(1) to INFREQ SIM PRACTICE(9) Ideal value= 2
(6)	3	2	6	4	8	5	SIMPLE TASK(1) to COMPLEX TASK(9) Ideal value= 4

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	RCIC OP AFTER LOSP	NI OP AFTER LOSP	SP COOL AFTER MSRV OPEN	MAN RX SCRAM ON HI SP TEMP	HI SP TEMP/PREV HPCI FAIL	RHR S/D COOL/RESTORE	Performance Shaping Factor
(1)	0.80	0.40	0.80	0.00	1.00	0.80	LITTLE TIME AVAILABLE to MORE TIME AVAILABLE relative importance = 0.24
(2)	0.25	0.00	0.50	1.00	0.50	0.75	HIGHER CONSEQUENCE to LOWER CONSEQUENCE relative importance = 0.19
(4)	0.33	0.33	1.00	0.67	0.00	1.00	MORE DIAGNOSIS REQ to LESS DIAGNOSIS REQ relative importance = 0.20
(5)	1.00	1.00	0.50	0.83	0.00	0.83	FREQ SIM PRACTICE to INFREQ SIM PRACTICE relative importance = 0.19
(6)	0.75	0.50	0.50	1.00	0.00	0.75	SIMPLE TASK to COMPLEX TASK relative importance = 0.19

The following scales are no longer in use for the reasons given below

(3) MORE SIM TRAINING to LESS SIM TRAINING
  You cancelled this scale during the session
END OF SUMMARY.

GROUP: 3
TASKS: LEVEL A, SUBSET 2

SUMMARY OF SESSION G3LAS SO FAR:

Current order of assessment of Likelihood of Success of actions
from best to worst
(Success Likelihood Indices are given in brackets)

LOCA W/NO SP COOL (0.72) (BEST

LOCA/INIT RHRSW (0.68)

HI SP TEMP/HPCI FAILURE (0.54)

TOTAL LOSS OF IA (0.26)

STATION BLACKOUT/ROOM COOL (0.13) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	HI SP TEMP/HPCI FAILURE	STATION BLACKOUT/ROOM COOL	LOCA/INIT RHRSW	LOCA W/NO SP COOL	TOTAL LOSS OF IA	Performance Shaping Factor
(1)	9	1	9	9	1	LOCAL OPERATION(1) to REMOTE OPERATION(9) Ideal value= 9
(2)	3	1	8	9	3	CUE AMBIGUITY(1) to CUE CLARITY(9) Ideal value= 9
(3)	9	8	3	3	9	FREQ SIM PRACTICE(1) to INFREQ SIM PRACTICE(9) Ideal value= 2
(5)	5	3	2	2	4	HIGH CONSEQUENCE(1) to LOW CONSEQUENCE(9) Ideal value= 6

* The following information shows your assessment of Likelihood of Success

1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	HI SP TEMP/HPCI FAILURE	STATION BLACKOUT/ROOM COOL	LOCA/INIT RHRSW	LOCA W/NO SP COOL	TOTAL LOSS OF IA	Performance Shaping Factor
(1)	1.00	0.00	1.00	1.00	0.00	LOCAL OPERATION to REMOTE OPERATION relative importance = 0.19
(2)	0.25	0.00	0.88	1.00	0.25	CUE AMBIGUITY to CUE CLARITY relative importance = 0.29
(3)	0.00	0.17	1.00	1.00	0.00	FREQ SIM PRACTICE to INFREQ SIM PRACTICE relative importance = 0.24
(5)	1.00	0.33	0.00	0.00	0.67	HIGH CONSEQUENCE to LOW CONSEQUENCE relative importance = 0.28

The following scales are no longer in use for the reasons given below

(4) DETAILED PROC. to GENERAL PROC.
You cancelled this scale during the session

END OF SUMMARY.

GROUP: 3
TASKS: LEVEL A, SUBSET 3

SUMMARY OF SESSION G3LAS3 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

MANUAL ADS AFTER LOSP (0.76) (BEST

ATWS/MANUAL INSERT (0.70)

ATWS WITH SLC (0.61)

STN BLCKOUT/RESTORE ONSITE (0.00) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	MANUAL ADS AFTER LOSP	ATWS WITH SLC	STN BLCKOUT/RESTORE ONSITE	ATWS/MANUAL INSERT	Performance Shaping Factor
(2)	2	4	9	5	MODERATE COMPLEXITY(1) to HIGH COMPLEXITY(9) Ideal value= 1
(4)	8	3	1	4	HIGH STRESS(1) to MODERATE STRESS(9) Ideal value= 9
(6)	5	3	7	2	ADEQUATE TRAINING(1) to INADEQUATE TRAINING(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	MANUAL ADS AFTER LOSP	ATWS WITH SLC	STN BLCKOUT/RESTORE ONSITE	ATWS/MANUAL INSERT	Performance Shaping Factor
(2)	1.00	0.71	0.00	0.57	MODERATE COMPLEXITY to HIGH COMPLEXITY relative importance = 0.29
(4)	1.00	0.29	0.00	0.43	HIGH STRESS to MODERATE STRESS relative importance = 0.32
(6)	0.40	0.80	0.00	1.00	ADEQUATE TRAINING to INADEQUATE TRAINING relative importance = 0.40

The following scales are no longer in use for the reasons given below

(1) HIGH STRESS to MODERATE STRESS
This scale was cancelled because of its
similarity with scale number 3
(3) HIGH CONSEQUENCE to LOW CONSEQUENCE
This scale was cancelled because of its
similarity with scale number 1
(5) FREQ SIM PRACTICE to INFREQ SIM PRACTICE
You cancelled this scale after trying to rate the actions on it.
END OF SUMMARY.

GROUP: 3
TASKS: LEVEL β, SUBSET 1

SUMMARY OF SESSION G3LBS1 SO FAR:

Current order of assessment of Likelihood of Success of actions
from best to worst
(Success Likelihood Indices are given in brackets)

METER OUT OF RANGE (1.00) (BEST

WRONG SWTCH W/MIMIC (0.79)

WRONG METER W/MIMIC (0.79)

JAMMED METER (0.21)

WRONG METER/FUNCT GROUP (0.16) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	WRONG SWTCH W/MIMIC	WRONG METER W/MIMIC	WRONG METER/FUNCT GROUP	JAMMED METER	METER OUT OF RANGE	Performance Shaping Factor
(1)	5	5	2	4	7	BAD DESIGN(1) to GOOD DESIGN(9) Ideal value= 9
(2)	2	2	4	5	2	HIGH FAMILIARITY(1) to LOW FAMILIARITY(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	WRONG SWITCH W/MIMIC	WRONG METER W/MIMIC	WRONG METER/FUNCT GROUP	JAMMED METER	METER OUT OF RANGE	Performance Shaping Factor
(1)	0.60	0.60	0.00	0.40	1.00	BAD DESIGN to GOOD DESIGN relative importance = 0.53
(2)	1.00	1.00	0.33	0.00	1.00	HIGH FAMILIARITY to LOW FAMILIARITY relative importance = 0.47

END OF SUMMARY.

GROUP: 3
TASKS: LEVEL B, SUBSET 2

SUMMARY OF SESSION G3LBS2 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

SIMILAR METERS W/LABELS (0.77) (BEST

CHART RECORDER/NORMAL BAND (0.69)

SIMILAR SWTCHS W/LABELS (0.69)

SIMILAR SWTCHS/FUNCTIONS (0.61)

WRONG CIRCUIT BREAKER (0.47)

CHART RECORDER/NO NORMAL BANDS (0.23) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	SIMILAR SWTCHS W/LABELS	SIMILAR SWTCHS/FUNCTIONS	SIMILAR METERS W/LABELS	WRONG CIRCUIT BREAKER	CHART RECORDER/NORMAL BAND	CHART RECORDER/NO NORMAL BANDS	Performance Shaping Factor
(1)	3	2	4	6	7	8	LOW DEMAND COGNITION(1) to MODERATE DEMAND COG(9) Ideal value= 5
(2)	6	5	7	6	4	8	GOOD DESIGN(1) to BAD DESIGN(9) Ideal value= 1
(3)	3	3	3	8	5	5	HIGH FAMILIARITY(1) to LOW FAMILIARITY(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	SIMILAR SWTCHS W/LABELS	SIMILAR SWTCHS/FUNCTIONS	SIMILAR METERS W/LABELS	WRONG CIRCUIT BREAKER	CHART RECORDER/NORMAL BAND	CHART RECORDER/NO NORMAL BANDS	Performance Shaping Factor
(1)	0.50	0.00	1.00	1.00	0.50	0.00	LOW DEMAND COGNITION to MODERATE DEMAND COG relative importance = 0.31
(2)	0.50	0.75	0.25	0.50	1.00	0.00	GOOD DESIGN to BAD DESIGN relative importance = 0.31
(3)	1.00	1.00	1.00	0.00	0.60	0.60	HIGH FAMILIARITY to LOW FAMILIARITY relative importance = 0.38

END OF SUMMARY.

GROUP: 3
TASKS: LEVEL β, SUBSET 3

SUMMARY OF SESSION G3LBS3 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

5 ANNUNCIATORS (1.00) (BEST

10 ANNUNCIATORS (0.45)

SIMILAR LOCAL VALVES (0.45)

SINGLE LOCAL VALVE (0.33) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	SIMILAR LOCAL VALVES	SINGLE LOCAL VALVE	5 ANNUNCIATORS	10 ANNUNCIATORS	Performance Shaping Factor
(1)	4	2	1	5	LOW IQ(1) to HIGH IQ(9) Ideal value= 1
(2)	4	7	2	3	HIGH FAMILIARITY(1) to LOW FAMILIARITY(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	SIMILAR LOCAL VALVES	SINGLE LOCAL VALVE	5 ANNUNCIATORS	10 ANNUNCIATORS	Performance Shaping Factor
(1)	0.25	0.75	1.00	0.00	LOW IQ to HIGH IQ relative importance = 0.44
(2)	0.60	0.00	1.00	0.80	HIGH FAMILIARITY to LOW FAMILIARITY relative importance = 0.56

END OF SUMMARY.

GROUP: 4
TASKS: LEVEL A , SUBSET 1

SUMMARY OF SESSION 1 SO FAR:

Current order of assessment of Likelihood of Success of ,
from best to worst
(Success Likelihood Indices are given in brackets)

MAN RS HI SP TEMP (0.69) (BEST

SP COOL MSRV OPN (0.66)

action"" (0.00)

actions (0.00)

RCIC OP LOSP (0.00)

NI OP LOSP (0.00) (WORST

Ratings of , on the scales you are currently using

Rating scale number	action""	actions	RCIC OP LOSP	NI OP LOSP	SP COOL MSRV OPN	MAN RS HI SP TEMP	Performance Shaping Factor
(4)	1	0	1	9	9	9	MORE COMPLEX(1) to LESS COMPLEX(9) Ideal value=-.4102643
(5)	1	0	1	9	1	1	FUEL DAMAGE(1) to NO DAMAGE(9) Ideal value=-.4925183
(6)	0	1	5	8	2	4	ADEQUATE FEEDBACK(1) to INADEQUATE FEEDBACK(9) Ideal value= 0

* The following information shows your assessment of Likelihood of Success
  for the , under consideration

  1.00 represents the best and 0.00 represents the
  worst on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end
  of this summary.

Rating scale number	action""	actions	RCIC OP LOSP	NI OP LOSP	SP COOL MSRV OPN	MAN RS HI SP TEMP	Performance Shaping Factor
(4)	9.00	4.00	0.00	1.00	1.00	1.00	MORE COMPLEX to LESS COMPLEX relative importance = 0.00
(5)	1.00	4.25	1.00	0.00	1.00	1.00	FUEL DAMAGE to NO DAMAGE relative importance = 0.00
(6)	0.00	0.00	0.00	0.00	0.00	0.00	ADEQUATE FEEDBACK to INADEQUATE FEEDBACK relative importance = 0.00

The following scales are no longer in use for the reasons given below

(1) HI SP TEMP PREV HPCI FAIL to RHR SD COOL REST
You cancelled this scale after trying to rate the , on it.
(2) PROCEDURE NOT AVAILABLE to PROCEDURE AVAILABLE
You cancelled this scale after trying to rate the , on it.
(3) HIGHER STRESS to LOWER STRESS
You cancelled this scale after trying to rate the , on it.
END OF SUMMARY.

GROUP: 4
TASKS: LEVEL A, SUBSET 2

SUMMARY OF SESSION G4LAS2 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

HIGH SP TEMP (0.59) (BEST

LOCA RHRSW (0.55)

LOCA NO SP COOL (0.54)

LOSS OF IA (0.41)

STATION BLACKOUT (0.40) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	HIGH SP TEMP	STATION BLACKOUT	LOCA RHRSW	LOCA NO SP COOL	LOSS OF IA	Performance Shaping Factor
(1)	9	9	1	1	8	ADEQUATE PROCEDURE(1) to INADEQUATE PROCEDURE(9) — Ideal value= 1
(2)	7	5	3	6	8	LESS REQUIRED TRAINING(1) to MORE REQUIRED TRAINING(9) — Ideal value= 9
(3)	2	4	4	8	2	HIGH FAMILIARITY(1) to LOW FAMILIARITY(9) — Ideal value= 2
(4)	8	4	3	2	2	HIGH PUCKER FACTOR(1) to LOW PUCKER FACTOR(9) — Ideal value= 8
(5)	7	4	1	2	9	CLEAR-CUT SITUATION(1) to AMBIGUOUS SITUATION(9) — Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	HIGH SP TEMP	STATION BLACKOUT	LOCA RHRSW	LOCA NO SP COOL	LOSS OF IA	Performance Shaping Factor
(1)	0.00	0.00	1.00	1.00	0.13	ADEQUATE PROCEDURE to INADEQUATE PROCEDURE
						relative importance = 0.19
(2)	0.80	0.40	0.00	0.60	1.00	LESS REQUIRED TRAINING to MORE REQUIRED TRAINING
						relative importance = 0.24
(3)	1.00	0.67	0.67	0.00	1.00	HIGH FAMILIARITY to LOW FAMILIARITY
						relative importance = 0.15
(4)	1.00	0.33	0.17	0.00	0.00	HIGH PUCKER FACTOR to LOW PUCKER FACTOR
						relative importance = 0.19
(5)	0.25	0.63	1.00	0.88	0.00	CLEAR-CUT SITUATION to AMBIGUOUS SITUATION
						relative importance = 0.23

END OF SUMMARY.

GROUP 4
TASKS: LEVEL A, SUBSET 3

SUMMARY OF SESSION ADS AFTER LOSP SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

STATION BLKOUT RSTR DG (0.92) (BEST

ADS AFTER LOSP (0.83)

ATWS INSERT RODS (0.50)

ATWS-SLC (0.00) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	ADS AFTER LOSP	ATWS-SLC	STATION BLKOUT RSTR DG	ATWS INSERT RODS	Performance Shaping Factor
(1)	6	1	5	3	HIGH OPERATOR DISCIPLINE(1) to LOW OPERATOR DISCIPLINE(9) Ideal value= 9
(2)	7	2	9	6	HIGH ECONOMIC LOSS(1) to LOW ECONOMIC LOSS(9) Ideal value= 9

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	ADS AFTER LOSP	ATWS-SLC	STATION BLKOUT RSTR DG	ATWS INSERT RODS	Performance Shaping Factor
(1)	1.00	0.00	0.80	0.40	HIGH OPERATOR DISCIPLINE to LOW OPERATOR DISCIPLINE relative importance = 0.40
(2)	0.71	0.00	1.00	0.57	HIGH ECONOMIC LOSS to LOW ECONOMIC LOSS relative importance = 0.60

END OF SUMMARY.

SUMMARY OF SESSION G4LBS1 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

METER OUOT OF RNG (0.86) (BEST

WRONG SWITCH (0.85)

WRONG METER MIMIC (0.85)

WRONG METER GROUP (0.54)

JAMMED METER (0.46) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	WRONG SWITCH	WRONG METER MIMIC	WRONG METER GROUP	JAMMED METER	METER OUOT OF RNG	Performance Shaping Factor
(1)	7	7	3	9	9	DIFFICULT IDENTIFICATION(1) to EASY IDENTIFICATION(9) Ideal value= 9
(3)	1	1	1	9	3	EASY TO INTERPRET(1) to DIFFICULT TO INTERPRET(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	WRONG SWITCH	WRONG METER MIMIC	WRONG METER GROUP	JAMMED METER	METER OUT OF RNG	Performance Shaping Factor
(1)	0.67	0.67	0.00	1.00	1.00	DIFFICULT IDENTIFICATION to EASY IDENTIFICATION
						relative importance = 0.46
(3)	1.00	1.00	1.00	0.00	0.75	EASY TO INTERPRET to DIFFICULT TO INTERPRET
						relative importance = 0.54

The following scales are no longer in use for the reasons given below

(2) INTERPRETATION to RECOGNITION
You cancelled this scale during the session
END OF SUMMARY.

GROUP 4
TASKS: LEVEL B, SUBSET 2

SUMMARY OF SESSION G4LBS2 SO FAR:

Current order of assessment of Likelihood of Success of actions from best to worst
(Success Likelihood Indices are given in brackets)

CR NBI (1.00) (BEST

CB LB (0.81)

SW LB (0.70)

MTR LB (0.70)

SW FN (0.38)

CR NO NBI (0.32) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	SW LB	SW FN	MTR LB	CB LB	CR NBI	CR NO NBI	Performance Shaping Factor
(1)	3	6	3	2	2	3	EASY TO IDENTIFY(1) to DIFFICULT TO IDENTIFTY(9) Ideal value= 1
(2)	2	2	2	2	1	4	EASY TO INTERPRET(1) to DIFFICULT TO INTERPRET(9) Ideal value= 1

* The following information shows your assessment of Likelihood of Success for the actions under consideration

  1.00 represents the best action and 0.00 represents the worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end of this summary.

Rating scale number	SW LB	SW FN	MTR LB	CB LB	CR NBI	CR NO NBI	Performance Shaping Factor
(1)	0.75	0.00	0.75	1.00	1.00	0.75	EASY TO IDENTIFY to DIFFICULT TO IDENTIFTY relative importance = 0.43
(2)	0.67	0.67	0.67	0.67	1.00	0.00	EASY TO INTERPRET to DIFFICULT TO INTERPRET relative importance = 0.57

END OF SUMMARY.

SUMMARY OF SESSION G4LBS3 SO FAR:

GROUP 4
TASKS: LEVEL B, SUBSET 3

Current order of assessment of Likelihood of Success of actions
from best to worst
(Success Likelihood Indices are given in brackets)

VALVE W/ID (1.00) (BEST

5ANNUN ALARM (0.62)

VALVE NO ID (0.44)

10 ANNUN ALARMED (0.35) (WORST

Ratings of actions on the scales you are currently using

Rating scale number	VALVE NO ID	VALVE W/ID	5ANNUN. ALARM	10 ANNUN ALARMED	Performance Shaping Factor
(1)	1	9	8	6	DIFFICULT TO IDENTIFY(1) to EASY TO IDENTIFY(9) Ideal value= 9
(2)	9	9	4	2	HIGH WORKLOAD(1) to LOW WORKLOAD(9) Ideal value= 9

* The following information shows your assessment of Likelihood of Success
  for the actions under consideration

  1.00 represents the best action and 0.00 represents the
  worst action on each Performance Shaping Factor.

* If you wish to change anything, you may do so at the end
  of this summary.

Rating scale number	VALVE NO ID	VALVE W/ID	5ANNUN ALARM	10 ANNUN ALARMED	Performance Shaping Factor
(1)	0.00	1.00	0.88	0.63	DIFFICULT TO IDENTIFY to EASY TO IDENTIFY relative importance = 0.56
(2)	1.00	1.00	0.29	0.00	HIGH WORKLOAD to LOW WORKLOAD relative importance = 0.44

END OF SUMMARY.

ATTACHMENT B-2

2. TECHNICAL RECORDER'S NOTES OF CONTENTS OF GROUP DISCUSSIONS DURING SLIM-MAUD SESSIONS

1. INTRODUCTION

A key distinguishing feature of SLIM-MAUD, vis-a-vis other subjective techniques for estimating human errors, is the emphasis placed upon its application in a group context. Group discussion and elaborations in SLIM-MAUD applications provide useful information for obtaining a more comprehensive understanding of the processes used by the group in arriving at its estimates. Accordingly, users of SLIM-MAUD are encouraged to torque-record SLIM-MAUD sessions in order to take advantage of the information on group process. Typically, two types of discussions and elaboration will be evident: task elaboration and PSF (Performance Shaping Factors) elaborations - each having the goal of arriving at a clear and shared understanding of tasks and the factors influencing their completion.

1.1 Elaborations in the SLIM-MAUD Evaluation

As noted above, considerable elaboration takes place within groups in a typical SLIM-MAUD session. As part of the SLIM-MAUD evaluation, the audiotapes of the SLIM-MAUD sessions were transcribed and summarized. Task descriptions can be found in Attachment B-1.

1.2 Task Elaborations by Group

Group 1:

Level A Tasks, Subset 1

Tasks were elaborated in some detail when describing how they should score on the PSF: Low complexity - High complexity.

Level B Tasks, Subset 2

In task 2 ("operator chooses wrong switch - grouped according to functions"), the group found it necessary to assume that all switches were also clearly identified by labels.

Task 10 (re: circuit breakers) was elaborated through a discussion of how circuit breakers are arranged in a PWR station.

Group 2:

Level A Tasks, Subset 1

The group decided to rule out consideration of information about current plant state in making its judgments, since in all cases the task description gave insufficient information on this.

Level B tasks, Subset 2

The group noted that only the level A descriptions permitted judgments about "real" tasks, as contextual information was given only in the task descriptions provided at that level - i.e., contextual information was not available for the B level tasks.

Group 3:

Level A Tasks, Subset 1

There was much discussion by the operators within the group concerning the type of Boiling water Reactor being assessed in RCIC OPERATION AFTER LOSP, as this vitally affected the rating of this task on the time dimension. The group agreed to assume a BWR4.

In general, in elaborating task consequences, scenarios were used to explore tasks in order to decide upon appropriate ratings.

Discussion ensued about task 7: HIGH SP TEMPERATURE - PREVENT POTENTIAL HPCI FAILURE. It was pointed out that one needs to know how high the temperature is in order to rate this task on the time pressure PSF. The problem of how to rate actions involving suppression pool cooling on simple-complex scale depends on the version of BWR in which the tasks were carried out. The group discussed this and again decided to assume BWR4.

Level A Tasks, Subset 3

In discussing task 9: ATWS - INITIATE SLC, it was noted that "this is the only task where it states that the operator does consult the procedure. In the others, you have to infer that the operator consults the procedure."

Level B Tasks, Subset 1

The group discussed the impossibility of bringing in motivational factors (differential effects of stress, sloppiness, etc) because no contextual information was supplied in the level B task descriptions which would allow it to do so.

Level B Tasks, Subset 3

The description of the task 15 involving ten annunciators was discussed at some length: there was a problem of variability of annunciator location from plant to plant and the fact that sometimes they were grouped together, sometimes not. Therefore, the task description was considered to contain insufficient information. It was noted that in some reactors it is possible to turn off all annunciators by turning one switch (e.g. Browns Ferry). The group finally decided to assume the operator sees the lights and acknowledges the first 9 annunciators.

Group 4:

Level B Tasks, Subset 2

In task 10, where auxiliary operator selects wrong circuit breaker, the group discussed the importance of a responsible operator having to instruct the auxiliary operator before going outside the control room.

1.3 PSF Elaborations

Group 1:

Level A Tasks, Subset 1

The PSF long-time available to short-time available (to carry out the task) was interpreted by the group in terms of the effects of time pressure on task performance.

There was an extensive discussion about what makes a task complex and much material about the possible context of the tasks was bought in here. However, the PSF low complexity to high complexity was finally subsummed under a second PSF describing the degree of training required. The group arrived at this decision after the problem of the non-independence of these two PSFs was raised by MAUD.

Level A Tasks, Subset 2

There was some confusion over what the PFS low cue to high cue meant; was it "many cues were available to the operator in the situation, anyway" or "that the operator needed to recognize many cues in order to perform successfully"?

In addition, a long discussion ensued (with similar confusion) about the PSF poor procedures to good procedures. Did this mean "procedures required" or "procedures available?" Finally, the group chose to define this PSF in terms of the degree to which procedures were needed to carry out the task (this was therefore not the same as group 2's PSF mandated (required) by procedures).

Level B Tasks, Subset 1

The PSF characterizing the cognitive complexity of the tasks was defined as complexity in carrying out the task, rather than complexity of the physical attributes of situation (e.g., number of switches).

The group noted that level A PSFs don't really apply to B level tasks (i.e., in general, tasks at different levels involve different PSF domains).

Level B Tasks, Subset 2

The PSF Low visual acuity to high visual acuity, was defined to be the amount of visual acuity one required to succeed in the task.

Level B Tasks, Subset 3

The group decided that the meaning of the PSF on degree of environmental impact referred to how much the immediate environment would be distracting to the operator (relevant in deciding between the annunciator tasks).

Group 2

Level A Tasks, Subset 2

There was discussion of the PSF requires training and the PSF requires diagnosis after MAUD had spotted non-independence of these 2 PSFs. The group decided that training required was subsumed under requires diagnosis.

Level B Tasks, Subset 1

After discussion, the group decided that the PSF high training required meant high training and experience required.

Level B Tasks, Subset 2

The group noted that the PSF operator experience also included degree of operator training required.

Group 3:

Level A Tasks, Subset 1

There was much discussion within the group about the correct position of the ideal point on some PSFs - e.g., on the PSF higher consequence - lower consequence: "Operator will be sloppy if the task is of low consequence but highly stressed if task is of high consequence" - so the group decided to rate the ideal point on this PSF at 3 fairly high consequence = best performance.

Concerning the PSF more simulator training - less simulator training, the group was confused about whether they should be rating the amount of simulation training required to perform the task well, or the amount of simulation training likely to have been had by the operators performing the task. A problem arose when the group decided to put the ideal point at the pole of the PSF where more simulator training was required - "i.e., how much required training would be ideal for the operators to have had." In view of this, the group reinterpreted the simulator training PSF as the amount of training which would have been provided in the training schedule for the task.

Because of the difficulty with the simulator training PSF, the group was not satisfied with the first MAUD result. So they re-entered MAUD, and deleted the old "simulator training" PSF. After discussion they replaced it with a PSF describing "frequency of simulator practice" (and re-rated the tasks on this new PSF).

Then the group decided to address another aspect of training, for they now recognized that previous discussion of the PSF task complexity was irrelevant" and now wanted to discuss tasks in terms of their "intrinsic complexity" (i.e., concerning what one has to do in carrying them out). Note that the group placed the ideal point in the middle of the scale on the complexity PSF, as "Operator will then have to think more carefully about a task that scores at this point on the scale than he would about a completely simple task."

Level A Tasks, Subset 2

There was a discussion about the meaning of the PSF frequent simulator practice. It was noted that it is important that the simulator practice should be adequate as well as frequent. MAUD identified similarity of ratings on the PSF's cue clarity and simulator training. The group then tried to find situations where these two PSFs were not similar in meaning and succeeding in doing so (just a few situations). So the group decided that while these two PSFs may be correlated due to plant design errors they both contributed independently to the likelihood of success, and so should both be left in.

There was a long discussion, as with the subset 1 tasks, about where to put the ideal point on the consequences scale. The group finally decided upon six as the ideal point; note this represents a shift from the ideal point of 3 on this PSF for subset 1 tasks.

Level A Tasks, Subset 3

MAUD found the PSF concerning degree of stress correlated with the PSF concerning degree of complexity but the group decided these 2 PSFs did not have the same meaning. They chose to use a PSF scaled from high stress to moderate stress in order to be able to place the ideal point at 9 on the scale - at the end describing moderate stress.

Level B Tasks, Subset 1

The group revised the results of their first MAUD solution, because on the PSF concerning familiarity they had confused familiarity with the device (e.g., meter) with familiarity with the task.

Level B Tasks, Subset 2

The group confirmed PSF familiarity means familiarity with the task (as in subset 1).

Level B Tasks, Subset 3

The meaning of the PSF pole High IQ was discussed and the group decided it meant "high IQ required to carry out the task successfully."

Group 4:

Level A Tasks, Subset 1

With regard to PSF higher stress to lower stress, there was a discussion of whether the operators stress level depends on if the operator knows he has "screwed up." There was a long discussion (and disagreement) on whether stress comes from lack of feedback. A member of the group with operations experience claimed that feedback entirely determines stress level, but a human factors member strongly disagreed. An operator in the group claimed that PSF weights would depend on the background of operator "I can tell you who came from economically driven plants - looking for short cuts - and who came from the navy - totally safety driven".

At the end of the session most members of the group made the following statement about the results: "results are fine: absolutely made sense, but I would never have guessed while going through the procedure." Then, in reviewing the MAUD audit trail, a typical comment was: "This is a very interesting finding for PRA - it says that it is very important for operators to know about consequences."

Level A Tasks, Subset 2

With regard to the PSF required training, there was a careful discussion of what was meant by "training" with the group agreeing that they should rate how much training operators would be required to have for successful performance of the task.

There was a discussion of how familiarity is different from training, and it may relate to experience (some confusion between guaranteed familiarity with a task and needed familiarity to perform a task sucessfully). There was also much discussion about the PSF concerning levels of stress, sources of stress, of what aspects of the situation an operator should be aware, and how this awareness would affect his stress level.

Level A Tasks, Subset 3

The PSF pole high operator discipline was described in contrast to panic, i.e., it implied that high operator discipline would be required to carry out the task successfully.

The group wished to use only 2 PSFs, as "every pathway to failure had to do with failure through worry - of economic loss, or of not being disciplined enough to handle the situation."

MAUD discovered that these two PSFs were highly correlated, but the group retained them both as independent contributors to the likelihood of success.

Level B Tasks, Subset 1

The group discussed the impossibility of bringing in motivational factors (e.g., differential effects of stress, sloppiness, etc.) because no contextual information was supplied which would allow them to do so.

The group rejected the first SLI's provided by MAUD when it was found that they had interpreted familiarity as familiarity with device rather than familiarity with the task (e.g., "spot a familiar meter was jammed" vs " meter was familiar but spotting that it was jammed was not"). This led to an ambiguous description of the PSF concerning interpretation/ recognition. So the group redefined this PSF as easy to interpret to difficult to interpret to disambiguate the way they thought about the familiarity issue, re-rated tasks and the ideal point on this PSF, and were then satisfied with the result produced by MAUD Level B tasks, Subset 2.

There was a discussion of a PSF concerning the amount of instruction that would be required (see "task elaboration") but they did not include any PSF to deal with this issue as they did not consider that, on balance, having to give instructions would increase failure probability on the task - "it might even help."

APPENDIX C

SUCCESS LIKELIHOOD INDICES AND HUMAN ERROR PROBABILITY ESTIMATES

C.1.0 INTRODUCTION

SLIM-MAUD is a subjective judgment technique for eliciting "decomposed" estimates of human reliability from experts familiar with nuclear power plant tasks. Specifically, SLIM-MAUD is a computer-based methodology used by a group of experts on a personal computer. The procedures elicit and quantify experts' judgments about specific factors affecting the likelihood of a task being performed successfully in a nuclear power plant. This methodology can be contrasted with techniques relying on the "wholistic" estimates of human reliability from experts, such as direct numerical estimation.

The experts who participated in the evaluation of SLIM-MAUD used SLIM-MAUD to produce Success Likelihood Indices (SLIs) for 15 Level A tasks and 15 Level B tasks. Then, they used direct numerical estimation procedures to wholistically estimate human error probabilities (HEPs) for the same 30 tasks. These results are presented in this appendix.

C.1.1 Success Likelihood Indices

SLIs are produced for a set of tasks which are assessed simultaneously during a group session. SLIs range from 0.00 to 1.00 and represent the relative likelihood of success for tasks assessed within a single session. Table C.1 presents the SLIs produced by each group for each subset of Level A tasks and Level B tasks. SLIs are based on the relative ratings of tasks on PSFs; therefore it is meaningless to compare SLIs for tasks assessed in different sessions since different PSFs are generally used across sessions. SLIs can be converted to HEPs and compared using a calibration program described in Appendix A, Section 7. However, for the purposes of evaluating the practicality, acceptability, and usefulness of SLIM-MAUD it was not necessary to convert the SLIs to HEPs. Therefore, only the SLIs for these tasks are presented in Table C.1.

C.1.2 Human Error Probabilites

Following the SLIM-MAUD session, each subject matter expert used direct estimation techniques to wholistically estimate exact HEPs, and upper and lower uncertainty bounds for the HEPs for the 15 Level A tasks and 15 Level B tasks. These estimates were done individually.

Table C.2 presents the HEPs and uncertainty bounds produced by direct estimation techniques for each of the tasks. The tasks correspond to the same Level A and Level B tasks presented in Table C.1. Therefore, the task descriptions are not repeated.

Table C.1 Success Likelihood Indices (SLIs) Produced by Each Group*

Level A Tasks					Level B Tasks				
Subset 1 Tasks	Group 1	Group 2	Group 3	Group 4	Subset 1 Tasks	Group 1	Group 2	Group 3	Group 4
2	.24	.53	.63	.66	3	.82	1.0	.79	.85
3	.44	.61	.44	.69	7	.90	.85	.79	.85
4	.80	.46	.67	.82	8	.46	.46	.16	.54
5	.29	.50	.67	.60	12	.00	.00	.21	.46
7	.30	.57	.33	.34	19	1.0	.58	1.0	.86
14	.75	.59	.83	.90					
Subset 2 Tasks					Subset 2 Tasks				
6	.36	.56	.54	.59	1	.16	.83	.69	.70
8	.24	.07	.13	.40	2	.33	.90	.61	.38
12	.57	.70	.68	.55	9	.41	.65	.77	.70
13	.83	.68	.72	.54	10	.31	.21	.47	.81
15	.15	.37	.26	.41	17	1.0	.57	.69	1.0
					18	.28	.29	.23	.32
Subset 3 Tasks					Subset 3 Tasks				
1	.77	.91	.76	.83	5	0.0	.08	.45	.44
9	.40	.63	.61	.00	6	.92	.49	.33	1.0
10	.28	.15	.00	.92	14	1.0	.73	1.0	.62
11	.82	.44	.70	.50	15	.39	.45	.45	.35

*See Appendix B, Attachment B-2, Section A for task descriptions.

Table C.2 Human Error Probabilities (HEPs) for Level A Tasks Produced by Individual Participants*

	Group 1 Experts				Group 2 Experts			
Tasks	A	B	C	D	E	F	G	H
1	.002	.05	.02	.002	.08	.02	.001	.03
2	.002	.01	.05	.02	.16	.005	.05	.01
3	.0001	.07	.01	.01	.05	.01	.013	.015
4	.0005	.01	.01	.02	.1	.01	.025	.05
5	.005	.06	.1	.0007	.12	.001	.01	.01
6	.0001	.08	.003	.1	.14	.01	.2	.01
7	.0005	.06	.005	.02	.16	.01	.05	.05
8	.01	.2	.07	.05	.07	.005	.1	.02
9	.05	.03	.005	.002	.1	.002	.001	.01
10	.2	.2	.5	.02	.05	.02	.01	.01
11	.05	.005	.01	.02	.08	.005	.01	.01
12	.001	.025	.02	.02	.05	.002	.02	.015
13	.001	.02	.005	.003	.03	.02	.002	.15
14	.0001	.01	.02	.003	.08	.02	.002	.15
15	.005	.1	.05	.02	.005	.005	.05	.025

	Group 3 Experts				Group 4 Experts			
Tasks	I	J	K	L	M	N	O	P
1	.00002	.001	.001	.002	.0001	.002	.000005	.00001
2	.00005	.001	.001	.001	.0005	.001	.0001	.0001
3	.00001	.0001	.0001	.00001	.01	.003	.00005	.001
4	.001	.001	.005	.003	.0005	.01	.0005	.0002
5	.00002	.005	.0001	.0000005	.0005	.01	.0000001	.001
6	.003	.0005	.002	.2	.001	.005	.5	.001
7	.0001	.0005	.0005	.002	.0002	.005	.5	.001
8	.003	.001	.02	.1	.0005	.01	.0005	.0002
9	.000005	.00001	.001	.0000005	.005	.0005	.000001	.00005
10	.001	.001	.002	.001	.002	.0001	.02	.0005
11	.00005	.0005	.0005	.000001	.0001	.0001	.000001	.0002
12	.00005	.0002	.001	.003	.0002	.002	.00001	.00002
13	.000005	.0002	.001	.001	.0002	.001	.0000075	.000005
14	.000001	.00005	.0002	.00002	.0001	.001	.0005	.00005
15	.0005	.001	.001	.005	.01	.002	.01	.0002

*See Appendix B, Attachment B-1, Section A for task descriptions.

Table C.2 (Continued) Human Error Probabilities (HEPs) for Level B Tasks Produced by Individual Participants*

	Group 1 Experts				Group 2 Experts			
Tasks	A	B	C	D	E	F	G	H
1	.02	.06	.05	.01	.1	.05	.02	.1
2	.02	.06	.03	.003	.08	.01	.002	.15
3	.02	.01	.02	.002	.04	.02	.001	.1
5	.1	.33	.25	.02	.08	.01	.05	.2
6	.002	.05	.02	.005	.007	.005	.005	.15
7	.02	.007	.01	.003	.04	.02	.002	.1
8	.05	.03	.02	.01	.05	.05	.02	.15
9	.02	.03	.03	.002	.08	.05	.05	.1
10	.05	.01	.05	.02	.1	.05	.05	.5
12	.1	.1	.2	.2	.1	.05	.2	.05
14	.0002	.006	.005	.02	.001	.005	.01	.1
15	.05	.06	.05	.05	.08	.001	.1	.05
17	.005	.007	.02	.01	.02	.005	.02	.05
18	.01	.05	.05	.02	.12	.05	.05	.2
19	.005	.02	.01	.01	.02	.002	.02	.4

	Group 3 Experts				Group 4 Experts			
Tasks	I	J	K	L	M	N	O	P
1	.001	.002	.05	.0001	.005	.002	.00002	.0001
2	.005	.001	.01	.0005	.01	.005	.0005	.0005
3	.0005	.0005	.005	.00001	.001	.001	.00005	.001
5	.002	.002	.05	.0005	.1	.01	.0005	.001
6	.0001	.0001	.02	.000002	.0001	.0002	.000001	.00002
7	.0005	.0005	.005	.00001	.001	.001	.000075	.002
8	.003	.001	.005	.00005	.01	.001	.00001	.0001
9	.0005	.002	.05	.0002	.01	.002	.0001	.0001
10	.003	.002	.05	.0002	.002	.003	.0002	.0002
12	.005	.001	.05	.001	.1	.01	.000001	.0001
14	.000001	.00001	.005	.0000005	.0001	.000001	.05	.000005
15	.005	.0002	.01	.003	.01	.0002	.01	.0001
17	.001	.0005	.05	.000005	.006	.005	.00001	.00002
18	.003	.001	.05	.000005	.01	.005	.001	.0005
19	.0001	.0002	.05	.000005	.0002	.002	.000001	.00005

*See Appendix B, Attachment B-1, Section A for task descriptions.

Table C.3 presents the aggregated HEPs for the subject matter experts within each group, with the 95% Statistical Confidence Limits. The individual HEPs were obtained using the procedure described in Section B.2.5 of Appendix B. After the individual experts' HEP assessments were obtained, they were aggregated using formula 3 in Appendix B (Section B.2.6.3, page B-20).

C.2.0 CAUTIONS REGARDING USE OF THIS DATA

The purpose of collecting the human reliability estimates produced during this project was to evaluate the SLIM-MAUD methodology in relation to other methods of subjective judgment. Some support for the convergent validity and inter-group reliability for these estimates were found during the evaluation of the methodology; however, no means for establishing the criterion-related or predictive validity of these estimates were available. The project was not designed to produce data on human reliability at nuclear power plant tasks for general application during risk assessment. These estimates may be valid only for the specific BWR tasks assessed, and for the subject matter experts who participated in this project. Therefore, it is recommended that the methods reported in this volume be used, not the data.

Table C.3 Human Error Probabilities (HEPs) and 95% Statistical Confidence Limits Aggregated Across Experts Within Groups*

Level A		Aggregated	95% Statistical	Confidence Limits
Tasks	Group	HEPs	Lower	Upper
1	1	.008	.002	.04
	2	.02	.002	.10
	3	.0004	.00005	.004
	4	.00006	.000004	.0008
2	1	.01	.003	.05
	2	.03	.005	.12
	3	.0005	.0001	.002
	4	.0003	.00008	.0009
3	1	.005	.0003	.08
	2	.02	.009	.04
	3	.00003	.000008	.0001
	4	.001	.0001	.01
4	1	.006	.001	.03
	2	.03	.01	.09
	3	.001	.0002	.006
	4	.0008	.0002	.005
5	1	.02	.005	.10
	2	.01	.002	.08
	3	.00005	.000001	.002
	4	.0005	.000007	.03
6	1	.007	.0003	.19
	2	.04	.008	.22
	3	.005	.0004	.07
	4	.008	.0004	.16
7	1	.008	.001	.06
	2	.05	.01	.14
	3	.0005	.0001	.002
	4	.004	.00009	.15
8	1	.05	.02	.18
	2	.03	.008	.11
	3	.009	.001	.07
	4	.0006	.00007	.005

*See Appendix B, Attachment B-1, Section A for task descriptions.

Table C.3 (Continued) Human Error Probabilities (HEPs) and 95% Statistical Confidence Limits Aggregated Across Experts Within Groups*

Level A				
Tasks	Group	Aggregated HEPs	95% Statistical Confidence Limits Lower	Upper
9	1	.01	.002	.05
	2	.007	.0009	.05
	3	.00001	.0000005	.0003
	4	.0002	.000005	.007
10	1	.16	.04	.63
	2	.02	.008	.04
	3	.001	.0008	.002
	4	.001	.00009	.01
11	1	.02	.006	.04
	2	.01	.004	.05
	3	.00006	.000003	.001
	4	.00002	.000002	.0002
12	1	.01	.006	.04
	2	.01	.003	.05
	2	.01	.003	.05
	3	.0004	.00007	.003
	4	.0001	.00001	.001
13	1	.004	.001	.01
	2	.02	.004	.13
	3	.0002	.00001	.002
	4	.00005	.000004	.0007
14	1	.003	.0003	.03
	2	.03	.004	.19
	3	.00002	.000002	.0002
	4	.0002	.00006	.0009
15	1	.03	.007	.10
	2	.01	.004	.04
	3	.001	.0005	.003
	4	.003	.0004	.02

*See Appendix B, Attachment B-1, Section A for task descriptions.

Table C.3 (Continued) Human Error Probabilities (HEPs) and 95% Statistical Confidence Limits Aggregated Across Experts Within Groups*

Level B		Aggregated	95% Statistical	Confidence Limits
Tasks	Group	HEPs	Lower	Upper
1	1	.03	.01	.06
	2	.06	.03	.12
	3	.002	.0001	.02
	4	.0004	.00003	.005
2	1	.02	.005	.07
	2	.02	.003	.17
	3	.002	.0006	.009
	4	.002	.0004	.009
3	1	.009	.003	.03
	2	.02	.002	.13
	3	.0003	.00003	.004
	4	.0005	.0001	.002
4	1	.12	.03	.43
	2	.06	.02	.19
	3	.003	.0005	.02
	4	.005	.0004	.05
5	1	.01	.002	.04
	2	.01	.003	.07
	3	.0001	.000003	.004
	4	.00003	.000002	.006
6	1	.008	.004	.02
	2	.02	.004	.11
	3	.0003	.00003	.004
	4	.0006	.0001	.003
7	1	.02	.01	.05
	2	.05	.02	.12
	3	.0009	.0001	.007
	4	.0003	.00002	.006
8	1	.01	.004	.05
	2	.07	.05	.10
	3	.002	.0002	.02
	4	.003	.0002	.03

*See Appendix B, Attachment B-1, Section A for task descriptions.

Table C.3 (Continued) Human Error Probabilities (HEPs) and 95% Statistical Confidence Limits Aggregated Across Experts Within Groups*

Level B				
Tasks	Group	Aggregated HEPs	95% Statistical Lower	Confidence Limits Upper
9	1	.03	.01	.06
	2	.12	.04	.35
	3	.003	.0003	.03
	4	.0009	.0003	.003
10	1	.14	.10	.21
	2	.09	.04	.17
	3	.004	.0006	.03
	4	.0006	.000004	.10
11	1	.003	.0005	.02
	2	.009	.001	.06
	3	.00001	.0000002	.0008
	4	.00007	.0000006	.008
12	1	.05	.05	.06
	2	.03	.003	.23
	3	.002	.0004	.01
	4	.001	.0001	.01
13	1	.009	.005	.02
	2	.02	.007	.05
	3	.0006	.00001	.03
	4	.0003	.000009	.009
14	1	.03	.01	.06
	2	.09	.05	.18
	3	.0009	.00002	.04
	4	.002	.0006	.009
15	1	.01	.006	.02
	2	.03	.003	.23
	3	.0003	.000006	.01
	4	.00007	.000003	.002

*See Appendix B, Attachment B-1, Section A for task descriptions.

NRC FORM 335 (6-83) U.S. NUCLEAR REGULATORY COMMISSION	1 REPORT NUMBER (Assigned by TIDC, add Vol. No., if any)
BIBLIOGRAPHIC DATA SHEET	NUREG/CR-4016 BNL-NUREG-51828 Vol. II - RX,AN,RG, 1S
	2 Leave blank
3 TITLE AND SUBTITLE Application of SLIM-MAUD: A Test of an Interactive Computer-based Method for Organizing Expert Assessment of Human Performance and Reliability, Volume II: Appendices	4 RECIPIENT'S ACCESSION NUMBER
	5 DATE REPORT COMPLETED MONTH: May YEAR: 1985
6. AUTHOR(S) C.M. Spettell, E.A. Rosa, P.C. Humphreys and D.E. Embrey	7 DATE REPORT ISSUED MONTH: October YEAR: 1986
8. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code) Brookhaven National Laboratory Engineering Technology Division Department of Nuclear Energy Upton, NY 11973	9 PROJECT/TASK/WORK UNIT NUMBER
	10 FIN NUMBER A-3272
11 SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code) U.S. Nuclear Regulatory Commission Human Factors and Safeguards Branch Washington, DC 20555	12a TYPE OF REPORT Formal
	12b PERIOD COVERED (Inclusive dates)

13. SUPPLEMENTARY NOTES

14 ABSTRACT (200 words or less)

The U.S. Nuclear Regulatory Commission (NRC) has been conducting a multi-year research program to investigate different methods for using expert judgments to estimate human error probabilities (HEPs) in nuclear power plants. One of the methods investigated, derived from multi-attribute utility theory, is the Success Likelihood Index Methodology implemented through Multi-Attribute Utility Decomposition (SLIM-MAUD). This report describes a systematic test application of the SLIM-MAUD methodology. The test application is evaluated on the basis of three criteria: practicality, acceptability, and usefulness.

Volume I of this report presents an overview of SLIM-MAUD, describes the procedures followed in the test application, and provides a summary of the results obtained.

Volume II consists of technical appendices to support in detail the materials contained in Volume I, and the users' package of explicit procedures to be followed in implementing SLIM-MAUD.

The results obtained in the test application provide support for the application of SLIM-MAUD to a wide variety of applications requiring estimates of human errors.

15a KEY WORDS AND DOCUMENT ANALYSIS

Expert Judgment Technique
Human Reliability Analysis
Nuclear Power Plants
Human Error Probabilities
Success Likelihood Indices

15b. DESCRIPTORS

16 AVAILABILITY STATEMENT	17 SECURITY CLASSIFICATION (This report) Unclassified	18 NUMBER OF PAGES 241
	19 SECURITY CLASSIFICATION (This page) Unclassified	20 PRICE $