NUREG/CR-4188, Vol. 2
ORNL/TM-9570/V2

ornl

OAK RIDGE
NATIONAL
LABORATORY

MARTIN MARIETTA

Nuclear Power Plant Simulation Facility Evaluation Methodology: Technical Bases

K. R. Laughery, Jr.
R. J. Carter

This Work Performed for
U.S. Nuclear Regulatory Commission
Division of Risk Analysis and Operations
Under
DOE Interagency Agreement 40-550-75

OPERATED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Available from

Superintendent of Documents
U.S. Government Printing Office
Post Office Box 37082
Washington, D.C. 20013-7982

and

National Technical Information Service
Springfield, VA 22161

NUREG/CR-4188, Vol. 2
ORNL/TM-9570/V2
Distribution Category RX

Engineering Physics and Mathematics Division

NUCLEAR POWER PLANT SIMULATION FACILITY EVALUATION METHODOLOGY: TECHNICAL BASES

K. R. Laughery, Jr.*
R. J. Carter

NRC Monitor: J. C. Lowry
Human Factors and Safeguards Branch
Division of Risk Analysis and Operations

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

Project Manager: R. J. Carter
Oak Ridge National Laboratory

Manuscript Completed: November 1985
Date Published: January 1986

This Work Performed for
U.S. Nuclear Regulatory Commission

Division of Risk Analysis and Operations
Under
DOE Interagency Agreement 40-550-75

NRC FIN No. B0466

Work Performed Under
ORNL Subcontract No. 37X-739545V

*Micro Analysis and Design
9132 Thunderhead Drive
Boulder, Colorado 80302

Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
operated by
Martin Marietta Energy Systems, Inc.
for the
DEPARTMENT OF ENERGY
Under Contract No. DE-AC05-84OR21400

ABSTRACT

This report is Volume 2 of a two-part document which describes a research and development project conducted for the Nuclear Regulatory Commission (NRC). The purpose of the effort was to develop a methdology to evaluate the acceptability of nuclear power plant (NPP) simulation facilities for use in the simulator-based portion of NRC's operator licensing examination.

The proposed methodology is to be utilized during two phases of the simulation facility life-cycle, initial simulator acceptance and recurrent analysis. Initial evaluation is to be performed when a simulation facility is acquired, in the case of new simulators, and as soon as is practical for existing simulators. The first phase is aimed at ensuring that the simulation facility provides an accurate representation of the reference NPP. There are two components of initial simulator evaluation: fidelity assessment and a direct determination of the simulator's adequacy for operator testing (i.e., evaluation of operator/trainee performance). Recurrent evaluation is aimed at ensuring that the simulation facility continues to accurately represent the reference plant throughout the life of the simulator. This phase involves three components: monitoring reference plant changes, monitoring the simulator's hardware, and examining the data from actual plant transients as they occur.

Volume 2 describes the development of and technical bases for the evaluation methodology, including a discussion of the major issues, research base, and judgements/decisions made by the research team. Volume 1 is a set of guidelines/procedures to be used by individuals involved in the simulation facility evaluation process.

TABLE OF CONTENTS

	Page No.
ABSTRACT	iii
LIST OF TABLES	vii
FOREWORD	ix
1.0 INTRODUCTION	1
1.1 Purpose	1
1.2 Organization	1
2.0 UNDERLYING PHILOSOPHIES	2
3.0 THE HANDBOOK DEVELOPMENT PROCESS	3
3.1 Introduction	3
3.2 Steps in the Development Process	3
4.0 IMPORTANT ISSUES IN THE HANDBOOK	11
4.1 General Issues	11
4.1.1 Evaluating the Simulator as an Operator Licensing Vehicle Versus Evaluating the Simulator as a Training Device	11
4.1.2 "Whole" or "Part" Simulator Acceptance	13
4.1.3 Existing Simulators	14
4.1.4 NRC's Role in the Evaluation Process	15
4.2 Fidelity Issues	15
4.2.1 What Level of Fidelity Is Required	15
4.2.2 Baseline Data	17
4.2.3 Selection of Plant Parameters for Measurement	19
4.3 Operator Experimentation	20
4.3.1 The Ten Experimental Designs	20
4.3.2 Operator Performance Measurement	22
4.3.3 The Basis for Setting a 20% Difference Criterion	23
4.4 Recurrent Evaluation Issues	24
4.4.1 What Are the Underlying Assumptions Behind the Approach to Recurrent Evaluation	24
4.4.2 Measurement of Operators After Transients	25
4.5 Summary	26

TABLE OF CONTENTS

	Page No.
5.0 REFERENCES	27
APPENDIX A - Questionnaire Sent to NRC for the Objectives and Resource Limitations Analyses	30
APPENDIX B - Questionnaire Used for Discussions with Vendors	62
APPENDIX C - Contact Reports from Telephone Conversations with Industrial and Other Experts	69

LIST OF TABLES

Page No.

1. General Categories of Resources Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation 5

2. Personnel Resources Which Might Be Required for Nuclear Power Plant Test and Evaluation 6

3. Nuclear Power Plant Simulator Resources Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation 7

4. Nuclear Power Plant Control-Room Resources Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation 7

5. Special Equipment and Costs Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation 8

6. Categories of Time Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation 8

FOREWORD

This report describes one of three tasks performed during the second phase of a research program that was initiated in March of 1982 to provide a technical basis for the implementation of a systems approach to training (SAT) in the nuclear power industry. The work previously completed in this program is described in NUREG/CR-3414, "Evaluation of Training Programs and Entry Level Qualifications for Nuclear Power Plant Control Room Personnel Based on the Systems Approach to Training" (P. M. Haas, D. L. Selby, M. J. Hanley, and R. T. Mercer, 1983), and NUREG/CR-3523, "A Ranking Scheme for Making Decisions on the Relative Training Importance of Potential Nuclear Power Plant Malfunctions" (D. L. Selby and W. T. Hensley, 1984). First phase work included a review of taxonomies of human performance and the identification of likely performance-shaping factors to be considered in entry level and training requirements for nuclear power plant (NPP) control room personnel. In addition, a proposed structure was produced based on the systems approach to training which used guided rating forms to evaluate each element of training system design and a technique was developed to rank plant malfunctions for their importance in training.

This second phase of the research program, initiated in July, 1983, has been oriented toward the development of a series of tools to operationalize the SAT technical logic. The three research tasks in the second phase can be described as follows:

1. "Development of a methodology for identification of NPP control room operator characteristics." The goal of this task was the generation of a technique to link operator characteristics derived from descriptions of in-plant task behaviors to potential measurement instruments. The research resulted in an automated task analysis tool called TAPS (the task analysis profiling system) which outputs lists of skills, knowledges, abilities, and attitudes when plant job descriptions are typed in. In addition, TAPS lists potential measurement tests which can be used to measure operator abilities.

2. "Development of a methodology for evaluation of simulation facilities." The purpose of the methodology is to assess the acceptability of simulation facilities for use in the simulator-based portion of the operator licensing examination. The documentation of this task, which has been conducted by Micro Analysis and Design under subcontract to ORNL, is presented in a two volume set. Volume 1 is a handbook to be used by individuals involved in the simulation facility evaluation process. Volume 2 describes the development of

and the technical basis for the methodology, and is the subject of this report.

3. "Development of a methodology for training task selection." The objective of this task was to develop a method to aid NRC in the assessment of whether or not plant training developers are allocating the training of individual tasks to appropriate training methods and to provide NRC with a method to select tasks for training research. This task was addressed through the development of a computer-based task sorting (TSORT) program which provides a scientific basis for task-allocation decisions and at the same time reduces NRC manpower work loads through automation.

The documentation of Tasks 1 and 3 appears in the two volumes of NUREG/CR-3481. TAPS is described in Volume 2, "Nuclear Power Plant Personnel Qualifications and Training: TAPS - The Task Analysis Profiling System" (C. C. Jorgensen, 1985), and the training task selection methodology is discussed in Volume 1, "Nuclear Power Plant Personnel Qualifications and Training: TSORT - An Automated Technique to Assign Tasks to Training Strategies" (C. C. Jorgensen, 1984).

1.0 INTRODUCTION

1.1 Purpose

This report is intended to serve as a supplement to the Handbook for the Nuclear Power Plant Simulation Facility Evaluation Methodology (Ref. 1). Many of the procedures for simulation facility evaluation presented in this handbook were based on previous research in simulator evaluation and many were based entirely on the judgments of the authors.

This report is intended to provide three types of information on the judgment process. First, the points at which these judgments were made are clearly identified and defined. Second, the process for arriving at the judgments that were made is clarified. Third, to the extent possible, the reasoning behind these judgments is stated.

1.2 Organization

This report is organized into three additional sections. Section 2 briefly lists the philosophies that guided the handbook development. Section 3 describes the overall technical approach taken to develop the evaluation methodology. This section includes a discussion of the experimental designs identified, an analysis of the evaluation objectives chosen, and an assessment of resources used in the evaluation and constraints associated with the resources. Section 4 explores issues in the handbook that the authors believe may be controversial. These issues are discussed with respect to research and/or judgments that led to the proposed approach.

2.0 UNDERLYING PHILOSOPHIES

As the handbook was developed, several philosophical concepts guided the authors' efforts. They are presented here to provide the handbook user with some context for the handbook development.

First, the procedures must be practical as well as technically adequate. They must follow scientific rigor, while at the same time, being balanced with the constraints of applied field research and the actual industry setting in which they will be implemented.

Second, and somewhat counterbalancing the first, the first version of the simulation facility evaluation methodology should be conservative from the standpoint of safety.

Third, validation of the methodology proposed in the handbook including a trial application in the field would be later required. This validation would serve to ensure the feasibility, practicality, and technical adequacy of the approach.

Fourth, the product of this effort, the handbook, was intended to be procedurally rather than theoretically oriented.

Fifth, and finally, to the greatest extent possible, the handbook should encompass evaluation methodologies that have been developed and tested in other environments to minimize technical risk.

3.0 THE HANDBOOK DEVELOPMENT PROCESS

3.1 Introduction

The simulation facility evaluation procedures which are presented in the handbook were developed considering the environment in which a simulator evaluation is performed. The theory and practice of testing and evaluating training simulators are fairly well understood and documented (Ref. 2). To develop a practical methodology, one should first identify the experimental designs which might be used for conducting simulator evaluation and then define the objectives of the simulator evaluation process and the limitations which would guide the type of evaluation that could be performed. Only after this analysis should the procedures for simulator testing be developed.

The approach selected had been used in the development of simulator test and evaluation procedures for the U.S. Air Force. The main points of the test procedure development process are described in Reference 3. To summarize the key concepts, the development of any evaluation technique is analogous to a mathematical programming problem. In mathematical programming, a problem is described in terms of: 1) the objectives which one is seeking to maximize and 2) the constraints which may limit satisfaction of these objectives. Then, a solution to the problem is found that maximally satisfies the objectives while staying within the identified constraints (i.e., a feasible experimental design which maximally satisfies the evaluation objectives).

3.2 Steps in the Development Process

The first step towards the development of a simulator evaluation program was to determine the designs that were available for simulator evaluation. Using the mathematical programming analogy, this was equivalent to determining the solution space.

Based on the experience of the authors, a set of candidate methodologies was developed and subjectively evaluated with respect to their applicability to nuclear power plant simulation facility evaluation. A summary of the candidate methodologies and the extent to which each methodology was believed to address the evaluation objectives is described in Section 4.3.1 of this report.

Once the various evaluation designs were identified, the objectives of the simulator evaluation process had to be clearly determined. Though some general objectives of the test and evaluation process had been identified, more specific statements had to be defined. For example, should the process be evaluative (i.e., determining if the simulator provides adequate training) and/or prescriptive (i.e., if it does not provide training, finding out why). Within these fairly broad objectives, there were a host of potential subobjectives, not all of which might be of interest. The goal in this step of the process was to identify all possible subobjectives and then determine which of these were most and least important. This analysis allowed us to focus the simulator testing procedure appropriately.

The decisions as to which simulator test and evaluation objectives were most important was a management rather than a technical decision. Therefore, we attempted to determine these objectives through a questionnaire that was presented to the decision makers at the Nuclear Regulatory Commission (NRC). A copy of the questionnaire that was developed is presented in Appendix A. NRC determined that the primary objective of the simulator testing process should be to determine whether the simulator is an appropriate vehicle for evaluating operator performance. Restated in more familiar terms, is the simulator an acceptable tool for conducting operator licensing examinations? The handbook design effort, therefore, focused on developing a procedure which would answer this question.

The next step was to identify the resource limitations which would affect the simulation facility evaluation process. Any test and evaluation process requires resources which may be limited. The resources could be personnel, material, or less tangible resources such as time or technology. For any evaluation procedure to be useful, one must consider the resources the procedure will require in terms of the resources which are likely to be available. A list of the five categories of potentially constrained resources is included in Table 1 and more detailed lists of subcategories for each of the five categories are included in Tables 2 through 6.

A three-step process was used to determine the resource limitations. First, the major categories of resources which might be consumed in a simulation facility evaluation were defined. Second, this list was reviewed to determine which resource limitations could be explored directly and which needed a context. Direct examination would involve asking

Table 1

General Categories of Resources Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation

Personnel to conduct the tests, analyze the data, and document the results
The nuclear power plant simulator that is being evaluated
The nuclear power plant control room that is being simulated
Special equipment and costs for the testing process
Time to perform the testing process

Table 2

Personnel Resources Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation

Behavioral scientists
Simulator hardware experts
Simulator software experts
Simulator technicians (cannot reconfigure simulator but can make minor changes)
Simulator test technicians (to actually collect data)
Data entry specialists
Simulator instructors
Simulator operators
Trainees (for subjects)
Nuclear power plant operators

Table 3

Nuclear Power Plant Simulator Resources Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation

Simulator time for initial testing
Simulator time for recurrent testing
Simulator automatic data collection, analysis, and storage capabilities

Table 4

Nuclear Power Plant Control-Room Resources Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation

Access to the plant control room during the time that the simulator is being initially evaluated
Access to the plant control room on a regular basis after the simulator has been initially certified
Automated plant and control-room data collection, analysis, and storage capabilities during normal operations
Automated plant and control-room data collection, analysis, and storage capabilities during transients

Table 5

Special Equipment and Costs Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation

Videotaping of operator performance
Automatic operator performance data reduction tools
Travel costs for members of the "test team"
Computers for data analysis
Simulator vendor baseline data
Nuclear power plant baseline data
Nuclear power plant engineering-model analyses

Table 6

Categories of Time Which Might Be Required for Nuclear Power Plant Simulator Test and Evaluation

Time to prepare a test plan
Time to collect nuclear power plant baseline data
Time to conduct tests
Time to analyze data
Time to write reports
Time for recurrent evaluations

the designated experts "how much of this resource will be available?" Some resources, however, need a context before their limitations can be correctly determined. For example, the availability of computer resources can be made via direct examination. However, the availability of time on the training simulator for engineering-fidelity testing may be affected by the amount of simulator time required for other parts of the test and evaluation process. Therefore, this resource's availability should be evaluated in the context of other potential resource needs.

Third, data on resource limitations were to be collected primarily by NRC's review of a set of scenarios which were developed. A copy of the scenarios is also included in Appendix A. Each of the scenarios required arbitrarily differing amounts of each resource. The collective responses to these scenarios were to provide a reasonable picture of the limitations of resource categories and groups of categories. However, the scenarios were not reviewed by a large number of NRC representatives and, consequently, the thoroughness of our resource analysis is of some doubt.

Once the above steps to determine resource limitations were completed, an approach for the methodology was defined. This is fundamentally the same approach that is described in the handbook. The approach involves four basic steps: 1) planning the evaluation process, 2) collecting, analyzing, and interpreting data on the simulator's fidelity, 3) collecting, analyzing, and interpreting data on operators who use the simulator, and 4) monitoring the simulator throughout its lifetime to ensure that it is properly maintained and modified as needed.

During the methodology development process two trips were made to different vendors of nuclear steam supply systems and nuclear power plant simulators. The purposes of these visits were to discuss the methodology being proposed and to solicit responses, criticisms, and suggested alternative approaches from experts in industry. A questionnaire used during these discussions is exhibited in Appendix B. This step was consistent with the philosophy that the evaluation process must be realistic. Informal discussions with these representatives indicated that the methodology is, by and large, reasonable. To further this analysis, a number of telephone calls were made to other individuals in the industry (e.g., utilities and best-estimate engineering modelers) to obtain their reactions to verbal descriptions of portions of the evaluation approach. Again, the comments seemed to indicate that the methodology is feasible.

Different individuals expressed different concerns about the practicality of the methodology. However, no clear consensus emerged about a particular portion of the methodology which is inappropriate; criticisms were levied on different parts of the methodology rather uniformly. The contact reports which were prepared from these telephone calls are included in Appendix C.

4.0 IMPORTANT ISSUES IN THE HANDBOOK

As was stated earlier, the handbook intentionally does not include a discussion of the theoretical framework supporting the methodology. Rather, the handbook presents a set of procedures that would comprise a simulation facility evaluation. This approach was selected to minimize unnecessary information for handbook users.

As should be obvious to anyone reading the handbook, there were a significant number of judgments that went into the selection of the handbook's procedures and criteria. Some of these judgments were based on related research but many were based solely on the intuition of the authors.

Four major areas are discussed in the subsequent sections:

1. General issues
2. Fidelity measurement
3. Operator experimentation
4. Recurrent evaluation

The discussion on each area is subdivided into a series of issues. For each issue, a brief discussion is presented on 1) a statement of the issue, 2) what we know about the issue, 3) what we don't know about the issue, 4) justification for the selected option, and 5) what should be done to better address the issue. Items 2 (what we know about the issue) and 3 (what we don't know about the issue) describe the authors' knowledge of the issues rather than what may be universally known. These discussions should, therefore, not be perceived as statements on the general state of knowledge of the issues but, rather, statements of the <u>authors'</u> assumptions about the issues. By explicitly stating these assumptions, others with more accurate knowledge can better critique and modify the evaluation methodology prior to its implementation.

4.1 <u>General Issues</u>

4.1.1 Evaluating the Simulator as an Operator Licensing Vehicle vs. Evaluating the Simulator as a Training Device

<u>Statement of the issue</u> - A simulator can be used for two distinctly different purposes. First, a simulator may provide a means for training and practice on necessary skills for successful job performance. Second, a simulator

may provide a mechanism for evaluating operator abilities regardless of how the abilities were acquired, such as in a licensing examination. To evaluate a simulator, one must decide which of these is important.

What we know - The methods one might use for evaluating simulators for training vs. for licensing are distinctly different. These differences are described in detail in Reference 4. To evaluate a simulator for training, one must be concerned only with its ability to reduce operator training time on the actual equipment (e.g., in a control room). An appropriate measure of a simulator's training effectiveness is the number of training trials required on the actual equipment to reach a criterion level of performance. On the other hand, to use a simulator for testing, one must be able to highly correlate the operator's performance in the simulator to his performance on the actual equipment. In cases where the simulator test results in a license to operate, one must be confident that an operator who can operate the simulated system will be able to operate the actual system the first time. "Number of trials to criterion" is not an appropriate measure in this case. Literature on the subject indicates that one should expect a simulator that is to be used for licensing examinations to require higher fidelity than one which is to be used only to provide operator training (e.g., Ref. 5).

What we don't know - The greatest lack of knowledge is in the area of fidelity requirements. This will be extensively discussed in Section 4.2.2. Briefly stated, the data do not exist to permit the translation of simulator training or licensing utilization requirements into physical fidelity criteria. Since extensive transfer of training experiments have not yet been done and may be impractical in an actual nuclear power plant environment, much of our simulator testing must rely on some type of physical fidelity measurement.

Justification for the selected option - In the handbook, the procedures are aimed solely at evaluating the simulator's ability as an operator licensing tool. This was done under NRC direction.

What should be done to better address the issue - The discussion of what should be done will be reserved until the discussion of fidelity in Section 4.2.2.

4.1.2 "Whole" or "Part" Simulator Acceptance

Statement of the issue - A simulator probably will not function perfectly for all operator testing activities, nor will it fail for all operator testing activities. Rather, a simulator will frequently represent different tasks with varying degrees of success. Furthermore, it is suggested that in the simulation of nuclear power plants, the tasks of greatest concern, both with respect to simulation fidelity and essential operator skills and abilities, are those in which the operator must effectively deal with a plant transient. If an approach is adopted which either "accepts" or "rejects" the simulator, then, in many cases, one would be faced with the dilemma of accepting a simulator which is a poor representation of some transients or rejecting a simulator which provides a good representation of most aspects of nuclear power plant operation.

What we know - For most transients, the primary practice that the operator will receive is in the simulator. Is an inadequate representation better than none at all? Reference 6 suggests not. According to Reference 6, depending upon the degree of dissimilarity between the plant and simulator, an inadequate representation may be worse than none at all. Alternately, should one reject a nearly perfect (but not quite) simulator? Common sense suggests not. How, then, do we reconcile these two contradictory concerns?

What we don't know - Because of the proprietary nature of the methods of developing simulator software, one must be concerned about the underlying implications of one "poor transient simulation." Does this mean that 1) there is a poor representation in the simulator software of some aspect of plant behavior which we simply did not observe elsewhere or 2) is the simulator's behavior associated with software specific to that transient? Additionally, how bad is too bad or, restated positively, how close is close enough?

Justification for the selected option - The authors elected to accept the simulator on an operator-task-by-operator-task basis (an operator task in this context is, for example, responding to a particular transient). This seems to be the most reasonable path since it does not force "blanket" acceptance of a simulator which has known flaws but, alternately, it does not reject a nearly perfect simulator. Presumably, an operator licensing simulator examination could be conducted without testing an operator on all types of transients but, rather, on a representative subset of those transients which are simulated with high fidelity. The simulator will be certified based on the results of that subset.

What should be done to better address the issue - First, it would be fruitful to require that the simulator vendors clearly identify transient-specific segments of their software. This would facilitate the assessment of whether problems are specific to a segment of the simulation or represent general patterns. Second, the operator licensing process needs to be reviewed to assess what subset(s) of the list of possible transients the operator should be tested upon to ensure adequate overall performance.

Third, the implications should be considered of training (not testing) operator tasks which are not adequately represented in a simulator. The "something is better than nothing" philosophy which, in this case, may be incorrect, should be checked. It is not apparent how this would be facilitated within the current charter of the NRC simulation facility evaluation program.

4.1.3 Existing Simulators

Statement of the issue - Many nuclear power plant simulators will have been built prior to the implementation of this methodology or even prior to the general acceptance of simulator standards such as American National Standard (ANS) 3.5 (Ref. 7). Hence, these simulators may not meet the standards that are presented in the handbook. Should the utilities that demonstrated the foresight of obtaining a simulator early be penalized for not waiting?

What we know - Utilities procuring new simulators will be able to perform the required facility evaluations during acceptance testing of the simulator. Therefore, the cost of facility evaluation will be relatively greater for utilities who already have simulators. Also, the time to conduct the simulation facility evaluation of existing simulators may have to come from scheduled operator training time.

What we don't know - The extent of the problem is unknown and depends on a variety of factors including the age of the simulator, the age of the plant, the fidelity of simulation in older simulators, etc.

Justification for the selected option - The option selected is to treat all simulators equally, regardless of status. However, as this decision was made, it was hoped that NRC policy might permit other considerations to enter into decision making. For instance, simulators more than five-years old could be given a two-year hiatus from meeting these requirements, or the fidelity requirements could be made less demanding for some period.

What should be done to better address the issue - It is suggested that, as the simulation facility evaluation process is further developed and refined, the practical consequences of the policy treating all simulators equally be carefully considered.

4.1.4 NRC's Role in the Evaluation Process

Statement of the issue - Should NRC participate in this process as an active participant (e.g., actually conducting the tests), a reviewer (e.g., reviewing all test results), or an auditor (e.g., retaining the right to selectively review and examine any test records and data)?

What we know - NRC is time and resource constrained.

What we don't know - The ability of the utilities to conduct simulation facility evaluations is unclear. They are also time and resource constrained. It will be difficult for any utility to have all of the required expertise available on staff. How much can they rely on the simulator vendors or consultants? The less that the utilities can afford, the more that NRC must provide or the less rigorous that the testing can be.

Justification for the selected option - The authors defined the NRC's role as that of an auditor. This was a management decision made by NRC.

What should be done to better address the issue - Unless NRC chooses to change its level of participation, no further work is required.

4.2 Fidelity Issues

4.2.1 What Level of Fidelity Is Required?

Statement of the issue - A simulation facility's acceptability will be based largely on its engineering fidelity. In other words, if one compares the simulator's performance on several variables to a baseline (which is as close to actual plant performance as possible), the simulator's acceptability for licensing examinations depends on how closely the simulation compares to the baseline.

What we know - Most of what is known about the relationship of fidelity to operator training/licensing is with respect to transfer of training. Even here, very little is known. It is known that low fidelity can frequently permit high transfer of training. However, as was discussed in Section 4.1.1, one should expect that fidelity requirements for testing to be higher. Data presented in a Reference 4 support this need for higher fidelity.

What we don't know - In this case, far less is known than is necessary for making empirically-based statements regarding fidelity requirements. Two recent literature reviews (Refs. 8 and 9) indicate the lack of our understanding, even where the relatively easy training transfer issues are concerned. When attempting to make specific statements regarding the needed fidelity for operator licensing, our knowledge deficit is even greater.

Justification for the selected option - At the end of Section 5 of the handbook, fidelity criteria are specified. Requirements are defined for each of the critical operator displays and consider factors such as root mean square error, average error, maximum deviation, and the error t-score. Additionally, a method is defined for determining the overall acceptability of the simulator for a task depending upon the results from analysis of individual displays. These criteria are based almost entirely on ANS 3.5 standards. At this time, there is no better basis for this decision. It is the authors' belief that these criteria are probably more conservative than necessary. However, until better evidence is available, it is appropriate to err conservatively.

What should be done to better address the issue - There is probably no issue which will be more controversial than the setting of criteria for simulator fidelity. The acceptance of all simulators rests on these criteria. Because of the lack of supporting research, any decisions are difficult to justify unequivocally.

An obvious solution is a program of research to answer these questions. With several years of directed research using parametric simulator-to-simulator transfer studies, many of these questions could be answered. It will probably be necessary in the near term to incorporate some judgment in setting these criteria.

4.2.2 Baseline Data

Statement of the issue - To assess simulator fidelity, one must have a baseline against which to compare the simulation. The baseline could be as simple as an operator's experience or something as complex as the information from a safety parameter display system (SPDS) in a plant. What should the baseline be?

What we know - The four main sources of baseline data can be thought of as 1) operator opinions, 2) reference plant data, 3) similar-plant data, and 4) best-estimate engineering model data. We know several things about these types of baseline data as are summarized in the following paragraphs.

With respect to operator opinions, there is concern that individual operators may be relatively unreliable judges of fidelity, although a consensus of operators' judgments will be more likely to result in a better measure of fidelity. However, most operators of nuclear power plants see very few, if any, of the moderate-to-severe transients that are so important to licensing examinations. Therefore, operator opinions on moderate-to-severe transients are going to be comprised primarily of intuition, guesswork, and what they are taught in training. The authors believe that this is a poor foundation on which to construct a simulator baseline.

With respect to actual plant data, there is very little of it for existing plants because of the relative lack of sophisticated plant parameter data recording systems. Even when parameter recording systems do exist, the lack of moderate-to-severe transient data still poses a data quantity problem. In the future, plants will be equipped with SPDSs as per NUREG-0696 (Ref. 10) so these data may become more available.

With respect to similar-plant data, more data are available depending upon the degree of similarity required between the reference and similar plants. In the handbook, criteria for determining similarity are presented, although these were made solely on the basis of expert judgement.

With respect to best-estimate model data, there are a number of computer codes - TRAC, RELAP5, Modular Modeling System, Plant Modeling System Program, and others - that can be used as best-estimate computer models for nuclear power plants. However, many of these models are time consuming and expensive to set up for specific plants. Additionally, running a best-estimate model for different transients frequently involves extensive reparameterization of the models and consumes large amounts of computer time. However, they can provide extensive data on critical aspects of plant safety during moderate transients.

What we don't know - With respect to actual plant data for both reference and similar plants, one of the main unknown factors is how the guidance provided in NUREG-0696 will ultimately be implemented in nuclear power plants. Safety parameter display systems could prove to be an excellent source of future data if they are used in a way analogous to the flight recording boxes on commercial aircraft. At any rate, this will be of little use to the short-term problem of evaluating the current generation of simulation facilities. In the long run, the data from SPDSs could provide useful data for testing and improving the current generation of simulators as more transient data become available.

Another major question is what constitutes a "best-estimate" model and, perhaps more importantly, when does a simulator model become of best-estimate quality? Does it lie in the model's relationships to first principles, the size of the reactor-core submodels, or the time between variable reestimation?

Justification for the selected option - In the handbook, the simulation faclility examination process permits the use of reference plant, similar plant, and best-estimate model data in roughly that order of preference, although situations are defined in which best-estimate models would be preferred to similar plant data. This seems to take advantage of the best data available under all circumstances.

What should be done to better address the issue - First, the requirement for SPDSs should be refined so that data can be generated for simulator development and testing. All that may be needed is an orderly method for retrieving and saving the information after transients. Based upon the information we obtained from a vendor, this is entirely possible within the instrumentation required by NUREG-0696. However, the requirement is vague enough that whether the SPDSs are used for data generation depends on how the utilities interpret NUREG-0696.

Second, careful consideration should be given as to what constitiutes a best-estimate model. A rigorous set of guidelines should be developed to define when a model is a best-estimate model. These guidelines should be based on predictive and concurrent validation studies relating model predictions to actual plant performance data and/or experiments in scaled facilities. If these criteria are not well defined, there is a distinct possibility that, as the state-of-the-art in power plant modeling advances, the simulator users will be faced with a continually changing definition.

Finally, methods for constructing best-estimate models must be streamlined. The interest for this is already evidenced in theory in a panel discussion presented in Reference 11 and in practice by a variety of nuclear power plant modeling tools including those described in References 12, 13, and 14.

4.2.3 Selection of Plant Parameters for Measurement

Statement of the issue - For evaluating the simulator's representation of any plant transient, one cannot reasonably compare the simulation to the baseline for all plant parameters. Rather, one must select the subset of parameters which are important to the simulation of that task.

What we know - Through the conduct of task criticality analyses, we are developing an understanding of the critical aspects of human operator behavior (e.g., Refs. 15 and 16). In some cases, these analyses have led to an understanding of the parameters which are critical to some aspects of plant operation.

What we don't know - As of yet, the plant parameters which are critical to plant operation for most of the transients described in ANS 3.5 have not been defined.

Justification for the selected option - The approach selected is that, for each operator task, a maximum of ten operator displays are identified by experienced reactor operators and nuclear engineers/designers which represent the critical plant parameters. In this case, the extensive use of expert judgment is deemed to be a practical approach and consistent with the statements made earlier in this section regarding the value of operator opinions. The authors treat operators as experts in identifying the critical parameters but do not rely on operator opinions as the primary measure of the parameters themselves.

What should be done to better address the issue - It was suggested by a vendor that, via the expert-judgment approach presented in the handbook, we could develop a standard set of parameters for all ANS 3.5 operator tasks. This list would be different for boiling-water reactors than for pressurized-water reactors (PWR) and, possibly, different across PWR manufacturers.

4.3 Operator Experimentation

4.3.1 The Ten Experimental Designs

Statement of the issue - There are a number of experimental designs which might be used for conducting simulator evaluation. Which is (are) the most appropriate?

What we know - Probably the best statement of the potential experimental designs for simulator evaluation is presented in Reference 17. While many of these are more appropriate for evaluating a simulator's utility for training, they can all provide useful data for evaluating a simulator with respect to licensing examinations. Below are brief descriptions of each of the ten experimental designs:

1. Transfer of training model - One group of trainees, the experimental group, receives simulator training prior to further training or testing in the power plant. Another group of trainees, the control group, receives all its training in the power plant. This is analogous to predictive validation.

2. Self-control transfer model - The trainees are tested in the power plant, trained in the simulator, then tested in the power plant again. This is a variant of the transfer of training model.

3. Pre-existing control transfer model - This is similar to transfer of training except the control group is comprised of trainees who learned the job prior to the existence of a simulator.

4. Uncontrolled transfer model - This is a default situation, in which no control group is employed due to circumstantial difficulties.

5. Simulator-to-simulator transfer model - This is a special application of transfer of training in which trainees are trained in one simulator and tested in another.

6. Backward transfer model - Here, the experimental subjects are already qualified operators, and they are then tested in the simulator. This model is analogous to concurrent validation.

7. Simulator performance improvement model - Only the improvement of performance in the simulator is examined. This model is used when in-plant testing is impossible.

8. Simulator fidelity analysis model - Here, no trainees are involved. The physical correspondence of the simulator to the power plant is evaluated. This is closest to content and face validity.

9. Simulator training program analysis model - Used to obtain qualitative assessment, this model involves questioning whether appropriate and/or innovative techniques are used in the simulator. This represents another content validation approach.

10. Opinion survey model - This is another content validity model in which operators, instructors, training specialists, and trainees (face validity) are asked their opinions about the perceived value of various features of the simulator.

What we don't know - At this point, the feasibility of each of the ten designs is largely unknown.

Justification for the selected option - One of the authors' first efforts was to estimate the utility of each of the above ten designs with respect to the assessment of the simulator's training transfer and operator licensing capabilities. The following table shows the ratings assigned to each design (a value of "5" indicates that the design will provide excellent data; a value of "0" indicates that the design will provide no useful information).

Approach	Utility for Assessing Transfer	Utility for Assessing Licensing
Transfer of training	5	4
Self-control transfer	4	4
Pre-existing control transfer	4	4
Uncontrolled transfer	3	5
Simulator-to-simulator transfer	3	2
Backward transfer	4	3

Approach	Utility for Assessing Transfer	Utility for Assessing Licensing
Simulator performance improvement	0	2
Simulator fidelity analysis	2	3
Simulator training program analysis	0	1
Opinion survey	2	2

Once these evaluations were made and it had been determined that simulator licensing was the objective, the feasible designs had to be identified. This resources and constraints analysis is discussed in Section 3. Additionally, recent literature in the field was considered (Refs. 18, 19, 20, 21, 22, and 23) In the handbook, it is proposed that operator performance be measured by utilizing elements of the transfer of training model (with reactor operator trainees) and the backward transfer model (with experienced reactor operators) to supplement the primary evaluation approach, simulator fidelity analysis.

What should be done to better address the issue - The authors believe that the entire direct operator measurement portion of the simulator facility evaluation process should be scrutinized with respect to its worth. It became apparent during the preparation of the handbook that direct operator measurement would involve small sample sizes and, even then, only for a small percentage of operator activities, mostly normal procedures. Rather than incorporating some of the performance measurement and data collection systems presented in the handbook, something closer to an opinion survey may be worthwhile. While the data obtained from an opinion survey would not be as valid as data obtained from the transfer of training and backward transfer designs suggested, the additional costs of these transfer studies may not be warranted.

4.3.2 Operator Performance Measurement

Statement of the issue - The development of operator performance measurement techniques can range from the highly subjective, manual techniques to objective, automated performance measurement systems.

What we know - Generally, the more automated the system, the more costly it is to develop. Also, we know that objectivity may increase reliability of measurement, but not necessarily its validity. There are many instances of highly objective performance measurement systems which provide little or no information about important aspects of operator behavior (Ref. 24).

What we don't know - As was discussed in Section 4.3.1, the main question is whether the development and use of operator performance measurement is an appropriate activity to include in simulation facility evaluation. Also of concern is the routine acceptance by the industry of rigorous operator performance measurement, such as that described in Reference 25.

Justification for the selected option - The handbook permits the use of any type of operator performance measurement system that the utility prefers. The authors' objective is to encourage the use of any means of performance measurement which will facilitate the collection of operator performance data.

What should be done to better address the issue - With respect to simulation facility evaluation, the question is, again, should direct operator assessment be an integral component of the simulation facility evaluation. If so, efforts should be undertaken to develop valid and reliable nuclear power plant operator performance measurement systems. These methods should be developed to be used on-the-job. Additionally, institutional constraints to collecting operator performance data must be relaxed.

4.3.3 The Basis for Setting a 20% Difference Criterion

Statement of the issue - In the handbook, a criterion of no more than a 20% difference in operator performance in the simulator vs. operator performance in the plant is presented. Why 20%?

What we know - One of the most compelling psychological phenomenon is Miller's "7 plus or minus 2" (Ref. 26). This states, briefly, that humans are good absolute discriminators at something between five and nine levels of stimuli for many different types of stimuli, including evaluations of human performance.

What we don't know - The reliability or validity of the performance measures that might be used by the utilities during the simulation facility evaluation is unknown.

Justification for the selected option - The 20% criterion represents the lowest level of discriminability as per Miller's rule (i.e., five levels). This lax criterion seems appropriate given the state of performance measurement in the industry.

What should be done to better address the issue - If the level of discrimination that could be achieved by the performance measurement system was known, then one could estimate the appropriate criterion based on the level of acceptable type 1 error (saying the simulator is inadequate when it is adequate) and type 2 error (saying the simulator is adequate when it is not). This scientific approach could be developed around the principles of statistical quality control which have been used in other training system evaluations (e.g., Ref. 27). The requisite research to support this is the determination of the reliability, validity, and consequent discriminability of the performance measurement systems.

4.4 Recurrent Evaluation Issues

4.4.1 What Are the Underlying Assumptions Behind the Approach to Recurrent Evaluation?

Statement of the issue - Any ongoing simulator evaluation program should seek a balance between 1) ensuring that the simulator is performing adequately and 2) avoiding excessive testing requirements. How can this balance best be achieved?

What we know - One of the underlying assumptions behind the simulator evaluation program in the handbook is that one must ensure that the simulator hardware stays properly calibrated. One need not be concerned about simulator software once it has been initially validated. Thus, by constantly checking the hardware, one can be reasonably certain of maintaining simulator fidelity.

Simulator software must be updated to reflect plant modifications. When this occurs, some checking of the simulator software changes must be made to ensure that an acceptable level of fidelity was achieved.

What we don't know - As in the case of initial simulator fidelity assessment, we don't know the interactive effects of the submodels of the simulation software. In other words, if some segments of the simulator software were changed to reflect the plant modifications, what effects would this have on the validity of other simulation software? Obviously, interactions do exist between different plant components which are reflected in the best-estimate models, although how this is translated into simulator code is uncertain. This gives rise to another question "have we tested all of the second-order effects of software changes?"

Justification for the selected option - The option selected involves frequent testing of the simulator's analog-to-digital and digital-to-analog conversion hardware. The approach seeks to isolate the components of the simulator which might be affected by plant modifications and then test them using fidelity evaluation techniques similar to those used for initial simulation facility evaluation. Also, plant modifications are studied with respect to their effects on plant systems. These plant system effects are ultimately related to operator tasks for which simulator fidelity tests are to be reconducted. This approach seems to provide a solid basis for ensuring that the simulator operates properly while consuming a minimum of testing resources.

What should be done to better address the issue - The interactive effects of simulator software changes should be studied and characterized. However, the methods for constructing simulator software may differ too much among vendors for general characterizations to be made.

4.4.2 Measurement of Operators After Transients

Statement of the issue - If transients occur in the plant, operators who were involved in the transient should be brought to the simulator. Why?

What we know - An operator who has just experienced a transient should be an excellent source of information on the ability of the simulator to represent the transient. However, operators involved in moderate-to-severe transients will probably be very busy for several weeks after the transient assisting in the assessment of what happened in the plant.

What we don't know - It is unclear whether one will be able to reconstruct the simulation scenario quickly enough after the event to present it to the operators while they still have adequate recall of the important details of the transient. Additionally, operator activities in the interim may serve to increase the likelihood of their forgetting these details.

Justification for the selected option - It seems that collecting operator performance information after a transient is an excellent opportunity to collect actual plant transient data which are generally lacking. This opportunity is too good to ignore.

What should be done to better address the issue - At this point, try it and see if it works. If the scenarios can be quickly reconstructed on the simulator and the operators can be made available, the data collected should be very valuable for identifying simulator deficiencies.

4.5 Summary

In summary, there are a large number of unresolved technical issues which have impact upon the simulation facility evaluation methodology's power to distinguish appropriate vs. inappropriate simulators as well as the methodology's feasibility. The handbook has sought to develop a balance between effectiveness and practicality. However, it is essential that the many and varied assumptions which went into the development of the handbook be tested in a field trial in connection with an actual simulation facility prior to formal implementation of the procedures.

5.0 REFERENCES

1. K. R. Laughery, R. J. Carter, and P. M. Haas, "Nuclear Power Plant Simulation Facility Evaluation Methodology: Handbook," NUREG/CR-4188, Vol. 1 (ORNL/TM-9570/V1), 1985.

2. J. L. Ditizan and K. R. Laughery, "SAC Simulator Certification Literature Review," SETA TM 79-14.2/1, Calspan Corporation, Buffalo, NY, 1979.

3. K. R. Laughery, "An Application of Mathematical Programming Concepts to Behavioral Research Design," Applied Ergonomics, 13(4), 263-268, 1982.

4. K. R. Laughery and J. L. Ditizan, "SAC Simulator Certification Handbook," 6825-Q-1, Arvin Industries, Inc., Buffalo, NY, 1982.

5. K. R. Laughery, J. L. Ditizan, and G. M. Houtman, "Differences Between Transfer Effectiveness and Student Performance Evaluations on Simulators: Theory and Practice of Evaluations," Proceedings of the Third Interservice/Industry Training Equipment Conference, 214-222, 1981.

6. C. E. Osgood, "The Similarity Paradox in Human Learning: A Resolution," Psychological Review, 56, 132-143, 1949.

7. American Nuclear Society, "Nuclear Power Plant Simulators for Use in Operator Training," American National Standard 3.5, LaGrange Park, Il, 1981.

8. W. L. Rankin, P. A. Bolton, R. Shikiar, and L. M. Saari, "Simulator Fidelity and Training Effectiveness: A Comprehensive Bibliography with Selected Annotations," NUREG/CR-3726, 1984.

9. K. R. Laughery, "A Selected Review of the Recent (1979-1983) Behavioral Research Literature on Training Simulators," NUREG/CR-4206 (ORNL/TM-9445), 1985.

10. Nuclear Regulatory Commission, "Functional Criteria for Emergency Response Facilities: Final Report," NUREG-0696, 1981.

11. R. Bronson, R. Favreau, G. Marr, A. Pugh, and R. Swartz, "Desk Top Computers and Simulation: A Roundtable Discussion," Simulation, 44(1), 40-43, 1985.

12. T. Moore, "A Modular Code for Plant Dynamics," Electrical Power Research Institute Journal, 9(10), 36-38, 1984.

13. Central Electricity Generating Board, "Plant Modeling System Program," London, England, 1983.

14. Combustion Engineering, Inc., "C. E. Computer Software Catalog," Windsor, Ct, 1984.

15. W. G. Stillwell, D. A. Seaver, and J. P. Schwartz, "Expert Estimations of Human Error Probability in Nuclear Power Plant Operations: A Review of Probability Assessment and Scaling," NUREG/CR-2255, 1982.

16. D. A. Seaver and W. G. Stillwell, "Procedures for Using Expert Judgement to Estimate Human Error Probabilities in Nuclear Power Plant Operations," NUREG/CR-2743, 1983.

17. P. W. Caro, "Some Factors Influencing Air Force Simulator Training Effectiveness," TR-77-2, Human Resources Research Organization, Alexandria, Va, 1977.

18. D. K. Baer, C. C. Francis, J. M. Winders, D. W. Jones, P. M. Haas, L. S. Abbott, and R. L. Dietz, "Nuclear Power Plant Simulators: Their Use in Operator Training and Requalification," NUREG/CR-1482 (ORNL/NUREG/TM-395), 1980.

19. B. Kraje and N. Snidow, "Analytical Simulator Qualification Methodology," 2054-1, Electrical Power Research Institute, Palo Alto, Ca, 1983.

20. P. M. Haas, D. L. Selby, T. W. Kerlin, and L. Felkins, "Specification and Verification of Nuclear Power Plant Training Simulator Response Characteristics, Part II. Conclusions and Recommendations," NUREG/CR-2353, Vol. 2 (ORNL/TM-7985/P2), 1982.

21. W. J. Gray, J. A. Greshim, and C. E. Meyer, "Westinghouse Simulator Validation," paper presented at the Eastern Simulation Conference, 1985.

22. C. O. Brewer and D. Hoffman, "Use of Simulators for Training and Evaluation of Nuclear Power Plant Personnel," Simulators, J. S. Gardiner, Editor (The Society for Computer Simulation, La Jolla, Ca), 16(1), 125-127, 1985.

23. J. Wachtel, "New Initiatives in Nuclear Power Plant Simulation: Proposed New Regulations and Guidance from the U.S. NRC," paper presented at the Eastern Simulation Conference, 1985.

24. R. N. Isley and E. J. Miller, "The Role of Automated Training in Future Army Flight Simulators," FR-ED-76-27, Human Resources Research Organization, Alexandria, Va, 1976.

25. A. N. Beare, R. E. Doris, and L. H. Gray, "Simulator Experiments: Effects of Experience of Senior Reactor Operators and of Presence of a Shift Technical Advisor in a Boiling Water Reactor Control Room," ORNL/TM-9327, 1984.

26. G. A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Review, 63, 81-97, 1956.

27. S. L. Johnson, "Retention and Transfer of Training on a Procedural Task: Interaction of Task Strategy and Cognitive Style," DJ-6032-M-1, Calspan Corporation, Buffalo, NY, 1977.

Appendix A

QUESTIONNAIRE SENT TO NRC FOR THE OBJECTIVES AND RESOURCE LIMITATIONS ANALYSES

QUESTIONNAIRES FOR DETERMINING NUCLEAR POWER PLANT OPERATOR SIMULATION FACILITY EVALUATION OBJECTIVES AND CONSTRAINTS

BACKGROUND

ORNL is working with NRC towards the development of procedures for evaluating nuclear power plant simulation facilities. Over the past few years, there has been considerable concern expressed over the quality of simulation presented by many simulators as is evidenced by the increased interest in ANS 3.5 (Nuclear Power Plant Simulators for Use in Operator Training) and Regulatory Guide 1.149 (Nuclear Power Plant Simulation Facilities for Use in Operator License Examinations). It would be to the benefit of everyone if procedures for evaluating simulators could be developed to support these documents which would ensure that operators trained on them were able to perform their jobs effectively. Your response to this questionnaire would greatly help us develop the best procedures.

We are trying to develop procedures which are well aimed at the evaluation objectives. However, these simulator evaluation objectives are not sufficiently defined in any available documentation. Obviously, the better defined the objectives of the evaluation process, the more likely that the procedures we develop will satisfy the expectations of them.

In addition, we are striving to ensure that the evaluation procedures do not place unreasonable demands on either NRC or industry. It is of no use to anyone to develop elaborate evaluation procedures that require the collection of information which is, in fact, impossible to obtain.

PURPOSE OF THIS QUESTIONNAIRE

We need your assistance in identifying the most important objectives and resource limitations from the NRC perspective. We believe that the issues of simulator evaluation objectives are a matter of NRC policy and, therefore, best addressed by those of you who best understand NRC policy issues. Additionally, factors which may limit resources are probably best understood by those of you with past experience in NRC evaluation and monitoring. In a nutshell, we are looking for your advice and guidance prior to developing training simulator evaluation procedures so that what we develop is as close as possible to what you want.

IDENTIFICATION OF SIMULATOR EVALUATION OBJECTIVES

We have summarized the potential objectives of the process of training simulator evaluation into the following six goal statements. The first two objectives involve the evaluation of the simulation facility with respect to its ability to provide effective operator training and as a tool for evaluating operator proficiency (e.g., licensing examinations). The last four objectives aim at assessing the simulator's utility for other functions such as evaluating control-room layouts. Please review each of the following six potential objectives of operator training simulator evaluation:

Simulator Facility Evaluation

Objective 1--Can operators be tested in the simulator?
Simulator evaluation should determine whether operators can be tested in the simulator to assess their proficiency. In other words, if an operator performs well in the simulator, can we be reasonably sure that he will perform well in the plant?

Objective 2 - Is the simulator good for maintaining skills?
Simulator evaluation should determine whether the simulator training received by the already trained operator will ensure that the operator remains proficient in plant operation.

Other Simulator Evaluation Objectives

Objective 3 – Is the simulator good for initially training operators?

Simulator evaluation should determine whether the student with no prior plant experience but trained in a classroom can obtain the necessary skills through simulator training for operation of the power plant being simulated.

Objective 4 – Can the simulator be used to evaluate plant operation during transients?

Simulator evaluation should determine whether the simulator is a good mechanism for testing plant operation during transients when plant operation cannot be evaluated by other means.

Objective 5 – Is the simulator over-designed or under-designed?

Simulator evaluation should determine specific ways that the simulator is over- or under-designed so that future simulators can be redesigned appropriately.

Objective 6 – Can the simulator be used for other plant operator research activities?

Simulator evaluation should determine whether the simulator can be used for other experimentation involving plant operators such as to evaluate control-room designs of the actual power plant.

To determine the relative emphasis you think that we should give to each objective, please mark how important each objective is by making a mark at the appropriate location on each of the following six scales:

Objective 1 (Can operators be tested in the simulator?)

|___|___|___|___|___|___|___|

Not Important ... Extremely Important

Objective 2 (Is the simulator good for maintaining operator skills?)

| | | | | | | |

Not Important — Extremely Important

Objective 3 (Is the simulator good for initially training operators?)

| | | | | | | |

Not Important — Extremely Important

Objective 4 (Can the simulator be used for evaluating plant performance during transients?)

| | | | | | | |

Not Important — Extremely Important

Objective 5 (In what ways is the simulator over-designed or under-designed?)

| | | | | | | |

Not Important — Extremely Important

Objective 6 (Can the simulator be used for other plant operator research activities?)

| | | | | | | |

Not Important — Extremely Important

DETERMINATION OF RESOURCE LIMITATIONS FOR SIMULATOR EVALUATION

To evaluate the utilization of any one kind of resource which may be required for simulator evaluation, it is usually helpful to have a context in which that resource will be required along with other kinds of resources. For example, whether videotaping of operator performance in the simulator is reasonable would depend, at least, on for how long taping would be performed, how many operators, and what other operator performance data we would want to collect. Therefore, we believe that it is necessary to ask the questions about resources which would be required during simulator evaluation by creating several "scenarios" about the simulator evaluation process. Then, you can have a better concept of overall resource requirements and, consequently, you should be able to better determine when we have required more resources than we should.

We have prepared four scenarios of simulator evaluation for the purpose of assessing resource limitations which are likely to affect the process of simulator evaluation. Before going on to the scenarios, let us clarify some terms and phrases which will be used.

Throughout each of the scenarios, we will be referring to individuals with a variety of skills. To clarify what we mean by these individuals, the following is a list of titles used throughout these descriptions along with a brief definition of what skills we expect each individual to possess:

<u>Simulator hardware expert</u> - An individual with a background in the fundamentals of simulator hardware design. This individual will probably have an education or training in electrical engineering or a similar field.

<u>Simulator software expert</u> - An individual with a background in simulator software design. This individual will probably have an education or training in computer science, systems engineering, or a similar field.

Simulator test technician - An individual who has been trained in the use of simulators. This individual need not have any special type of education or training, as long as he can be trained to use the training simulator being evaluated and the associated test equipment.

Simulator instructor - An individual with experience in using simulators to train other individuals in the operation of nuclear power plants.

Reactor operator trainees - Individuals who have not been certified in the operation of the nuclear power plant for which the simulator is designed. Also, these individuals will never have been senior reactor operators at any nuclear power plant.

Senior reactor operators - Individuals who have been trained in the operation of the nuclear power plant for which the simulator is designed and who have been certified as senior reactor operators at this or another nuclear facility. If the simulator is designed for an existing plant, these individuals will have a minimum of one-year experience at the plant.

Behavioral scientist - An individual with training or experience in conducting experiments with humans. These individuals will have a degree in experimental psychology or a related field.

We also use the terms training simulator and engineering simulator. A training simulator is a simulator which is used primarily for training and testing operators. An engineering simulator is a simulator which is used to evaluate the engineering design of the plant. An engineering simulator could be used to evaluate control-room layout, coordinated operation of plant subsystems, or to answer other engineering questions. These terms are not mutually exclusive (i.e., training simulators are sometimes used as engineering simulators and vice versa).

Each of the following four scenarios has been separated into the following parts:

1. An NRC team which would be required for each initial simulator test as derived from the requirements of the evaluation process.
2. An industry team which would be required for each initial simulator test.
3. The type of information to be included into the test plan which would be prepared prior to evaluation (this test plan could be incorporated into acceptance test procedures or vice versa).
4. The steps which would be involved in initial simulator testing.
5. The types of information and data which would be collected during initial simulator testing.
6. The steps which would be involved in recurrent simulator rechecking.
7. The types of information and data which would be collected during recurrent simulator rechecking.
8. Levels of effort required to implement the scenario.

Throughout these scenarios, the distinction is made between <u>initial simulator testing</u> and <u>recurrent simulator rechecking.</u> Initial simulator testing is performed as the simulator is initially brought into the training program or, in the case of existing simulators, as soon as the evaluation procedures we are currently defining are finalized. Recurrent simulator rechecking is performed throughout the life of the simulator to ensure that the simulator maintains the level of performance observed during initial testing. As you read these, please keep in mind the applicability of the procedures to each of the following three types of simulator/plant combinations:

New plant / New simulator
Existing plant / New simulator
Existing plant / Existing simulator

We expect that the feasibility of these approaches will depend, to some extent, on which of these situations exist and we would like your input if a scenario may be reasonable for one but not others.

Please read each of the following four descriptions of how simulator evaluation might be conducted. As you read each one, please make notes in the space provided when you think that the simulator testing plan is in any way excessive. Please make comments liberally. We would prefer to find out now what is and is not feasible rather than after the procedures are developed. We also request that you focus your attention on the resources which are required in these descriptions rather than the procedures which are outlined, although any comments regarding procedural problems are also welcome.

SCENARIO 1

To initially evaluate each simulator, the following team of individuals will be put together by NRC:

1. One simulator hardware expert (eighth-time)
2. One simulator software expert (eighth-time)
3. One behavioral scientist (eighth-time)

The utility company that will use the simulator will be required to provide the following individuals:

1. One simulator hardware expert (full-time)
2. Two simulator software experts (full-time)
3. One simulator technician (full-time)
4. One simulator instrructor (full-time)
5. Four reactor operator trainees who will receive their initial training on the simulator
6. Four senior reactor operators who will receive their recurrent training on the simulator

Prior to commencement of testing, a formal test plan will be developed which will include the following:

1. A list of the source and location of all plant baseline data against which the simulator can be compared.
2. References to computer code for all engineering models which are available for the reactor being simulated.
3. A test schedule.
4. A list of utility participants in this study including the reactor operator trainees and senior reactor operators who will receive simulator training during testing.

The initial simulator certification will involve two steps:

1. An evaluation of the fidelity of the simulator
2. An evaluation of operator performance in the simulator and in the plant

Simulator fidelity will be measured by different methods depending upon whether the plant for which the simulator is designed is in operation or under development. Regardless of the plant's status, the selection of variables and tests will require an extensive analysis to determine the critical parameters and tests for simulator fidelity assessment. The development of these evaluation procedures will require approximately three months of the team's full-time effort.

If the simulator is being designed for an existing plant, then all normal operations plant data will be provided from actual plant data collection. If the plant has automated plant performance data collection capabilities, then data for abnormal or emergency operations will be provided for those transients which 1) are simulated on the simulator and 2) have occurred at the plant. If the transient has not occurred at the plant or the plant does not have automated plant performance data collection capabilities, then trasnsient data will be provided by exercising the engineering models which were identified in the test plan.

If the simulator is being designed for a new plant, then all plant operations data will be provided by exercising the engineering models. When the plant is brought on line, the engineering models' predictions for normal operations will be compared to actual plant data to ensure that the assumptions made from the engineering models were correct.

The industry team will be responsible for developing all data collection tools and collecting and analyzing the data. NRC will write the operating plant scenarios under which the data will be collected.

The simulator will pass the fidelity portion of the initial test if it meets the requirements identified by the system identification analysis. These criteria will vary from simulator to simulator.

The training effectiveness analysis during initial testing will be accomplished in two ways. First, the performance of the senior reactor operators on the simulator will be evaluated the first three times that they practice a normal operation on the simulator. If they are found to be making consistent errors, they will be questioned as to the source of the error. If their responses indicate that the simulator has some inaccuracies, this will be further explored. Additionally, the first time that the senior reactor operators perform an emergency procedure on the simulator which they have previously experienced in the plant, they will rate the simulator's accuracy of simulating that transient. This portion of the training effectiveness evaluation will be conducted for existing plants only.

Second, reactor operator trainees will receive normal training on the simulator. When they return to the plant, their performance during the first time that they perform each normal operation will be measured. If they make consistent errors, the simulator's fidelity will be reexamined on the portions of the simulator related to the tasks on which the operators consistently err.

All data collection and analysis will be performed by the industry team.

If the training effectiveness portion of the evaluation is passed without difficulty or if the difficulties are traced to simulator deficiencies which are then corrected, the simulator will be considered acceptable for operator training and testing.

Once the initial testing is completed, the simulator will be rechecked on a regular basis.

This rechecking will be accomplished by examining all analog-to-digital and digital-to-analog conversion points in the simulator. These conversions must be accurate within acceptable tolerance limits.

When there are plant changes which will affect the simulator, the simulator will not be used for training or testing any operator tasks which may be affected by the change. The determination of which tasks will be affected will be made by examining the systems affected by the plant changes and which tasks involve the use of those systems. The simulator will not be used for training these affected tasks until all of the initial fidelity tests are redefined (based on the plant changes) and reconducted on the simulator.

Based on this scenario, we can assume that initial certification will require approximately ten months.

END OF SCENARIO 1

SCENARIO 2

To initially evaluate each simulator the following team of individuals will be put together by NRC:

1. One simulator hardware expert (half-time)
2. One simulator software expert (quarter-time)
3. Two simulator test technicians (full-time)
4. One simulator instructor or past-instructor (quarter-time)

The utility that will use the simulator will be required to provide the following individuals:

1. One simulator hardware expert (half-time)
2. One simulator software expert (quarter-time)
3. Two simulator technicians (i.e.,individuals who cannot make engineering changes to the simulator but can run and make minor alterations in simulator operation (full-time)
4. One simulator instructor (full-time)
5. Two reactor operator trainees who will receive their initial training on the simulator
6. Three senior reactor operators who will receive their recurrent training on the simulator

Prior to the commencement of simulator testing, a formal test plan will be prepared by the utility which specifies the general nature of the evaluation to be conducted. This test plan will be approximately 15-20 pages long and will include the following:

1. A brief description of the actual plant performance data against which the simulator will be compared
2. A brief description of the plant engineering models against which simulator models will be compared

3. A test schedule
4. A list of individuals (NRC and utility) who will be involved

The initial simulator evaluation will involve two essential steps:

1. A subjective evaluation of the simulator's resemblance to plant operation for normal operating procedures (e.g., startup, generator synchronization) and frequently encountered transients (e.g., the initial phases of a turbine trip)
2. An evaluation of the fidelity of the simulator

The subjective evaluations of the simulator's resemblance to plant operation will be obtained from both the reactor operator trainees and the senior reactor operators, but in different ways. The trainees will evaluate the simulator in the actual plant after they have performed each task that they learned on the simulator for the first time. For example, the first time that the reactor operator trainee performs a generator synchronization in the plant (after simulator training), he will rate how similar the process of synchronization was in the actual plant compared to the simulator synchronization process. Likewise, the first time he encounters one of the frequently encountered transients, he will evaluate the simulator's accuracy of representation of that transient. The senior reactor operators, on the other hand, will evaluate the simulator as soon as they arrive for refresher training, both on normal and transient operations.

Simulator fidelity will be evaluated for both normal and transient operation. The actual plant baseline data will be used for all normal procedures and those transient operations for which data exists. No plant data collection will be required for this test.

For those transient operations for which adequate data do not exist, the engineering models will be run and their predictions will be compared to those mathematical models used in the training simulator for all variables which are used in both the engineering and training simulator models. The basis of comparison will be a visual examination by NRC of a graph presenting the engineering vs. training models during the course of an event. The NRC software expert will decide if the models look reasonably similar.

Once the subjective and fidelity data outlined above have been collected, they will be summarized in a report which includes the following:

1. A summary of the testing process (i.e., the test plan previously described)
2. A list of the evaluations given by both the reactor operator trainees and the senior reactor operators for each normal operation and transient measured
3. Copies of all of the engineering vs. training model or plant baseline data vs. training model comparisons along with a rating of the acceptability of the training models

For every normal and transient operation evaluated, the simulator will either be accepted for testing or not. This acceptance will be based on an examination of the test results with respect to a predefined set of criteria (e.g., 2 out of 3 senior reactor operators indicated simulator acceptability and at least 75% of the variables compared in fidelity measurement were considered acceptable). Note that the simulator will be accepted on an operation by operation basis.

For those operations or transients not found to be acceptably simulated on the simulator, the utility company will be given the opportunity to modify the simulator to bring its performance into acceptable bounds.

Once the initial testing has been completed, the simulator will be rechecked on a regular basis. This rechecking will vary depending upon whether the need for a recheck is based on a plant modification or not.

If the need for a recheck is brought about by a plant change, the utility will be required to conduct the fidelity evaluation portion of the initial testing process for all systems and procedures which are affected by the change. Continuing acceptance of the simulator for operator evaluation of the affected procedures will be based upon the results of these tests.

For those systems that have not been altered by plant modifications, the utility will be responsible for checking all analog-to-digital or digital-to-analog conversion points throughout the simulator at least once every six months. If these conversions are off by more than a predefined percentage (e.g., 5 %), they will be readjusted or replaced immediately.

Based upon this scenario of simulator testing, we can expect the initial testing phase to require approximately four months.

END OF SCENARIO 2

SCENARIO 3

To initially evaluate each simulator, the following team of individuals will be put together by NRC:

1. One simulator hardware expert (quarter-time)
2. One simulator software expert (half-time)
3. One simulator test technician (full-time)
4. One simulator instructor or past-instructor (full-time)
5. One behavioral scientist (half-time)

The utility company that will use the simulator will be required to provide the following individuals:

1. One simulator hardware expert (quarter-time)
2. One simulator software expert (full-time)

3. Three simulator technicians (full-time)
4. One simulator instructor (full-time)
5. Eight reactor operator trainees who will receive their initial training on the simulator
6. Eight senior reactor operators who will receive their recurrent training on the simulator

Prior to commencement of the testing, a formal test plan will be prepared by the utility which will include the following:

1. A list of the source and location of all plant baseline data against which the simulator will be compared
2. Listings of computer code for all engineering models which were used in the development of mathematical models for the simulator
3. A test schedule
4. A list of NRC and utility officials who will be involved in the evaluation including the reactor operator trainees and senior reactor operators who will receive training on the simulator during testing

The initial simulator evaluation will involve two essential steps:

1. An evaluation of the fidelity of the simulator
2. A set of experiments which measure the performance of individuals trained in the simulator

Simulator fidelity will be measured by two different methods, depending upon whether we are testing simulation of normal operations or simulation of transients.

For normal operations, we will compare simulator performance to actual plant performance. The comparison will be made by collecting data on the values of operator displays during proper execution of the procedure in both the plant and the simulator. If these data do not exist in the plant baseline data, they will be obtained either by videotaping the control-room displays or, if possible, obtaining computer printouts of the display values. The data will be obtained in a similar manner in the simulator.

Once these data have been obtained for normal operations, data sampling techniques will be used for data reduction. For each operation, a correlation analysis will be performed comparing simulator and plant display values throughout each plant operation. If the correlation coefficients for all displays involved in the operation exceed a certain minimum value, the simulator will pass the fidelity portion of the simulator testing process. If not, the utility will be given the opportunity to modify the simulator as required.

For evaluating simulator operation during transients, the simulator's performance will be compared to the engineering models of the plant's performance. Again, the primary measures for each transient will be a correlation of the simulator vs. predicted operator display values at points throughout the transient. Therefore, only those portions of the engineering models which predict variables presented on operator displays will be used. The same data sampling and analysis techniques will be used as in normal operation (as described above).

Two sets of experiments will be conducted to directly evaluate the training that operators are receiving on the simulator.

The first experiments will be with the reactor operator trainees. During simulator training, the performance of these operators will be measured at each training session. Their performance will be evaluated in one of the two following ways:

1. By the use of a simulator performance measurement system which automatically collects and stores data on operator performance
2. By videotaping operator performance and then having it reviewed and evaluated at a later time by the NRC staff

After the reactor operator trainee has completed all simulator training, his performance during actual plant operation will be evaluated in one of the two ways described above. His simulator performance will then be compared with his performance in actual plant operation to assess the effectiveness of simulator training.

The second experiment will be conducted with senior reactor operators. As soon as the SROs arrive at the facility to receive simulation training, they will be tested on the simulator for all normal operating procedures and transients that they have encountered in the past two months. Any problems that they have in performing an operation will be recorded and, at the completion of the procedure, the operators will be asked what led them to make the errors. If it is determined that many of the SROs make the same mistakes for the same reasons, the simulator fidelity will once again be examined to determine the source of the problem.

At the completion of these experiments with reactor operator trainees and senior reactor operators, the simulator will be accepted for training all normal and transient operations which were found to have adequate fidelity, and for which the simulator was found to provide useful training.

Once the initial testing is completed, the simulator will be rechecked on a regular basis.

The fidelity will be rechecked by conducting one third of the initial fidelity tests on a rotating basis. These tests will be performed within fixed time periods so that the entire initial fidelity test is reperformed every three-time periods (e.g., tests every six months so that entire initial fidelity test is reperformed every eighteen months).

The training effectiveness will be monitored in two ways. First, every time a major transient occurs in the actual power plant, the operators who were on duty at the time of the transient will be brought to the simulator to "replay" the transient on the simulator. If these operators detect distinct differences between the plant's performance during the transient and the simulator's performance, the simulator will be restricted from training that transient or similar transients.

The second way of monitoring training effectiveness will be through a formal questioning of all senior reactor operators who receive refresher training on the simulator. Within the first 20 hours of simulator training, they will be asked to identify any elements of simulator operation which do not reflect actual plant operation. If a criticism is found consistently, the simulator will be restricted from training those affected operations until the deficiencies can be identified and corrected.

Based upon this scenario of simulator testing, we can expect the initial testing phase to require approximately nine months.

END OF SCENARIO 3

SCENARIO 4

To initially evaluate each simulator, the following team of individuals will be put together by NRC:

1. One simulator hardware expert (half-time)
2. One simulator software expert (full-time)
3. One simulator test technician (full-time)
4. One simulator instructor (full-time)
5. One behavioral scientist (quarter-time)

The utility that will use the simulator will be required to provide the following individuals:

1. One simulator hardware expert (half-time)
2. Two simulator software experts (full-time)
3. One simulator test technician (full-time)
4. Two simulator instructors (half-time)

Prior to the commencement of simulator evaluation, a formal test plan will be prepared by the utility which specifies the following:

1. The source and location of the plant baseline data which was used to develop the simulator mathematical models
2. The source and location of other plant baseline data which can be used for simulator model validation
3. The source and location of plant engineering models which were used to develop the simulator models
4. Other plant engineering models which can be used for simulator model validation
5. A test schedule
6. A list of NRC and utility individuals who will be involved in the test

The initial simulator evaluation will involve two essential steps:

1. An analysis of simulator fidelity
2. An analysis of the way that the simulator is to be used in operator training

Simulator fidelity will be evaluated by the following steps:

1. A review will be made of plant operation during all normal and transient operations to determine the critical variables that describe plant performance to the operators.
2. For normal and transient operations for which baseline data exists (that was not used for simulator model development), simulator performance with respect to those variables identified above will be compared to the baseline data.
3. For operations for which unique baseline data do not exist, simulator performance with respect to the above variables will be compared to the values predicted by all available plant engineering models.

All comparisons of simulator to plant baseline or engineering models will be made by computing a large number of "difference" scores (i.e., the difference between simulator and "correct" values for each variable). Then, trend analysis techniques and inferential statistics will be used to determine whether consistent and/or significant differences exist.

Additionally, to ensure adequate simulator fidelity, a senior reactor operator will conduct all normal and transient operations on the simulator. If the SRO feels that there are certain operations for which the simulator's performance is noticeably inaccurate, the fidelity data will once again be reviewed to determine if deficiencies exist.

Concurrent to the fidelity evaluation, an analysis will be conducted of how the simulator is to be used for training. The utility will provide documentation regarding how much training each type of operator will receive, on what specific operations, in what sequence, and how often. This information will then be reviewed by the NRC behavioral scientist and simulator instructor to determine whether they meet a predefined set of good instructional practices for simulator training. These predefined criteria will be available to the utility prior to simulator testing.

If the simulator is found to have adequate fidelity and the simulator training program is found to follow good instructional practices, the simulator will be accepted for training the qualified operations.

Once the simulator is accepted, the simulator will be rechecked annually. This annual recheck will be performed by two senior reactor operators who will perform all normal and transient simulator operations. They will use checklists to ensure consistency.

If the SROs detect deficiencies in the simulator's representation of a plant operation, the fidelity check conducted for that operation during initial testing will be reperformed. If deficiencies are identified, the simulator will be restricted from training the affected operations until the appropriate modifications are made.

Based upon this scenario of simulator evaluation, we can expect the initial testing phase to require approximately six months.

END OF SCENARIO 4

Let us now try to summarize the four scenarios:

Man-months required for initial certification

Person Type	Scenario Number 1	2	3	4
NRC Simulator hardware expert	0.8	2	2.25	3
NRC Simulator software expert	0.8	1	4.5	6
NRC Simulator test technician	0	4	9	6
NRC Simulator instructor or past-instructor	0	1	9	6
Industry Simulator hardware expert	10	2	2.25	3
Industry Simulator software expert	20	1	9	12
Industry Simulator test technician	10	4	27	6
Industry Simulator instructor	10	4	9	6
Behavioral scientist	0.8	0	0	3

Note: In all scenarios, senior reactor operator and reactor operator trainees are peforming little more than their normal training duties.

BRIEF DESCRIPTIONS OF THE SCENARIOS

Scenario 1

- The test plan will be fairly detailed including sources and locations of all data and computer code.

- Initial certification will rely largely on the measurement of the simulator's fidelity. Data will also be collected on the performance of trainees in the plant after they have received simulator training and on the performance of experienced operators in the simulator.

- Simulator rechecking will be accomplished by 1) carefully monitoring plant changes to ensure that the simulator is modified appropriately and 2) monitoring analog-to-digital and digital-to-analog conversion points.

- The initial certification process will require ten months.

Scenario 2

- The test plan will be very short, basically an "executive level" description of the test.

- Initial certification will involve two steps: 1) a subjective evaluation of the simulator for normal operating procedures and 2) an evaluation of simulator fidelity. No actual plant data collection wil be required for either step.

- Simulator rechecking will be performed by monitoring plant changes and by checking all analog-to-digital and digital-to-analog conversion points.

- The initial certification process will require about four months.

Scenario 3

- The test plan will be very detailed including specific sources and locations of all data and computer code.

- Initial certification will involve two steps: 1) an evaluation of simulator fidelity and 2) a set of experiments where operator performance is measured. Actual plant data will be collected for evaluating simulator fidelity of normal operations and for measuring the performance of individuals trained in the simulator.

- Simulator rechecking will be accomplished by reconducting portions of the initial certification fidelity tests on a regular basis.

- The initial certification process will require nine months.

Scenario 4

- The test plan will be fairly detailed indicating the sources and locations of important data to be used during testing.

- Initial certification will involve two steps; 1) an evaluation of simulator fidelity and 2) an evaluation of the ways that the simulator is used for training. No actual plant data will be collected for either step.

- Simulator rechecking will be accomplished annually by two senior reactor operators going through all simulator operations and denoting any deficiencies which are found.

- The initial certification process will require six months.

If you have any comments you would like to add, please provide them below.

Now, we would like you to go through the following list of personnel resources which could be required for initial simulator testing and evaluate how much of each resource we should expect to be available per simulator <u>for initial simulator evaluation only.</u>

PERSONNEL RESOURCES

Simulator hardware expertise from NRC

| | | | | | | | | |

00 hours — 2000 hours

Simulator software expertise from NRC

| | | | | | | | | |

00 hours — 2000 hours

Simulator technicians from NRC

| | | | | | | | | |

00 hours — 2000 hours

Simulator instructors from NRC

| | | | | | | | | |

00 hours — 2000 hours

Simulator hardware expertise from Industry

| | | | | | | | | |

00 hours — 2000 hours

Simulator software expertise from Industry

```
|    |    |    |    |    |    |    |    |    |
------------------------------------------------
00                                          2000
hours                                      hours
```

Simulator technicians from Industry

```
|    |    |    |    |    |    |    |    |    |
------------------------------------------------
00                                          2000
hours                                      hours
```

Simulator instructors from Industry

```
|    |    |    |    |    |    |    |    |    |
------------------------------------------------
00                                          2000
hours                                      hours
```

Simulator operators from Industry

```
|    |    |    |    |    |    |    |    |    |
------------------------------------------------
00                                          2000
hours                                      hours
```

Senior reactor operators from Industry who will be trained on the simulator

```
|    |    |    |    |    |    |    |    |    |
------------------------------------------------
00                                          2000
hours                                      hours
```

Reactor operator trainees from Industry who will be trained on the simulator

```
|    |    |    |    |    |    |    |    |    |
------------------------------------------------
00                                          2000
hours                                      hours
```

Now, we would like you to go through the following list of personnel resources which could be required for recurrent simulator checking and evaluate how much of each resource we should expect to be available <u>annually per simulator for recurrent simulator checking only.</u>

<u>PERSONNEL RESOURCES</u>

Simulator hardware expertise from NRC

| | | | | | | | | | |

00 hours per year — 200 hours per year

Simulator software expertise from NRC

| | | | | | | | | | |

00 hours per year — 200 hours per year

Simulator technicians from NRC

| | | | | | | | | | |

00 hours per year — 200 hours per year

Simulator instructors from NRC

| | | | | | | | | | |

00 hours per year — 200 hours per year

Simulator hardware expertise from Industry

| | | | | | | | | | |

00 hours per year — 200 hours per year

Simulator Software expertise from Industry

```
  |    |    |    |    |    |    |    |    |    |
  ----------------------------------------------
 00                                            200
hours                                          hours
per year                                      per year
```

Simulator technicians from Industry

```
  |    |    |    |    |    |    |    |    |    |
  ----------------------------------------------
 00                                            200
hours                                          hours
per year                                      per year
```

Simulator instructors from Industry

```
  |    |    |    |    |    |    |    |    |    |
  ----------------------------------------------
 00                                            200
hours                                          hours
per year                                      per year
```

Simulator operators from Industry

```
  |    |    |    |    |    |    |    |    |    |
  ----------------------------------------------
 00                                            200
hours                                          hours
per year                                      per year
```

Now please mark on the following scales the amount of non-personnel resources which should be available for initial certification or recurrent testing, as indicated.

SIMULATOR RESOURCES

Simulator time devoted to initial testing of the simulator (non-training time)

00 hours — 2000 hours

Simulator time devoted to recurrent testing of the simulator (i.e., annual rechecks)

00 hours per year — 2000 hours per year

Are simulators in the near future going to be built with automatic data collection devices (i.e., for performance testing)?

highly unlikely — highly likely

ACTUAL POWER PLANT CONTROL ROOM RESOURCES

Control-room time devoted to collecting data for initial testing of the simulator

00 hours — 2000 hours

Control-room time devoted to collecting data for recurrent testing of the simulator (i.e., annual rechecks)

| | | | | | | | | |
----

00 hours per year — 2000 hours per year

Are nuclear power plants in the near future going to be built with automatic data collection devices (e.g., for measurement of plant performance during transients)?

| | | | | | | | | |
----

highly unlikely — highly likely

Will it be possible to videotape operator performance in actual plant control rooms?

| | | | | | | | | |
----

highly unlikely — highly likely

Will it be possible to obtain plant performance data which was not used in developing the simulator mathematical models?

| | | | | | | | | |
----

highly unlikely — highly likely

We would like to thank you for your assistance in the development of simulator testing procedures. If you have any further questions or comments, please contact Mr. John Lowry at NRC-RES.

Once again, thank you for your help.

Appendix B

QUESTIONNAIRE USED FOR DISCUSSIONS WITH VENDORS

Where Does The Data Come From During Simulator Design?

1. For existing plants, what is the nature of the plant data you receive for simulating normal operations (e.g., format, content, amount)?

2. For existing plants, what is the nature of the plant data you use for simulating plant transient operations? What does the availability of these data depend upon?

3. Are engineering models used extensively in simulator code development?

4. Who provides the engineering model data?

5. When is this engineering model data provided?

6. Do you use all of the data (both engineering model and plant data) during simulator code development or is some saved for model validation?

7. Do you use expert opinion to design, modify, or validate the simulator code?

8. How are simulator deficiencies reported and/or repaired once the simulator is accepted by the customer?

Engineering Model-Specific Questions

1. What engineering models do you use for generating plant performance data for training simulator design?

2. Are these models general to PWRs or BWRs or are they plant specific?

3. Do you have integrated models of reactor operation or are the engineering models specific to plant subsystems?

4. If they are specific to subsystems, how are they integrated for a specific plant?

5. How are the models parameters set for a specific plant?

6. Are there variables in the "typical" plant engineering model which map directly into all operator displays?

7. Can the parameters be set for a plant so that different transients can be run without reparameterizing the entire model (except for the plant's starting condition and other parameters reflecting the nature of the transient)?

8. How much effort is required and by whom (person types and weeks) to set up an engineering model simulation of a transient which has never been simulated before?

9. How much effort is required and by whom to set up an engineering model simulation of a transient which has been simulated before but for which a slightly different plant starting configuration is desired?

10. How much computer cost is incurred in running the model for transient simulation (e.g., cost per second of transient simulated)?

11. How have your engineering models been validated and is there a continuing program of model validation?

12. Are these models proprietary?

Operator Training Questions

1. How are individuals that are being trained evaluated during training?

2. Would it be feasible to solicit operator opinions of the simulator during their training?

Miscellaneous Questions

1. Could an operator list the "ten most important displays" that he uses during the performance of a task?

2. Would the above list match a similar list compiled by a plant design engineer?

3. Will future nuclear power plants have automated plant performance measurement and recording capability?

Appendix C

CONTACT REPORTS FROM TELEPHONE CONVERSATIONS WITH INDUSTRIAL AND OTHER EXPERTS

CONTACT REPORT

Best-Estimate Engineering Modeler
8/3/84

Main Topic of Converstaion: The availability of best-estimate engineering models for the testing of nuclear power plant simulators.

The conversation began with a description of the general approach we were proposing for simulator evaluation. Then, we focused the conversation on our reliance on best-estimate engineering models for fidelity testing.

His first response was that "the approach was rational and probably the only way to do it." We then asked him to elaborate on his knowledge about the availability of best-estimate models.

He indicated that they have designed a best-estimate "simulator." It is a simulator since it works in better than real-time for most events. However it is not a realistic model for simulator vendors since it works in real-time on a CDC Cyber 7600 main-frame computer. Using this "generic" best-estimate analysis (BEA) model, specific PWR plants can be modeled by reparameterization of the model.

Reparameterization can be accomplished in approximately one man-month if the data are available. He indicated that sometimes these data are difficult to obtain for a specific plant. To remedy these data collection difficulties, NRC is sponsoring an effort being conducted by Technology Development of California entitled the "Nuclear Plant Data Bank." The apparent goal of this effort is to provide a "clearinghouse" for all plant performance information for all of the nuclear power plants in the United States. This data would be in a format which would allow the parameterization of their BEA models in one man-day rather than a month. Currently, this data bank is nearly empty. He suggested that we may want to reference the data bank specifically in our handbook as well as their "simulator" as sources of BEA data.

With respect to the handbook, we believe that this contact reconfirms our reliance on BEA data. Best-estimate models for PWRs are available to anyone who needs them if they can provide the needed plant parameters. Our feeling is that if they cannot provide the plant parameters, we should not accept the simulator.

CONTACT REPORT

Nuclear Power Plant Utility
8/13/84

Main Topic of Converstaion: The reasonableness of the approach proposed in the handbook.

The conversation began with a description of the general approach we were proposing for simulator evaluation. Then, we went through each step of the approach in greater detail and solicited his comments on the feasibility of the techniques.

With respect to fidelity evaluation, he indicated that the proposed approach is generally what they do now. They use BEA data during simulator design and, when possible, they used actual plant data. The concept of having to collect data for every operator "task" (which we defined for him) seemed reasonable.

With respect to the actual operator data collection, he indicated that our approach sounded reasonable. He indicated that ORNL was already collecting "that kind of data" and that it was not an unreasonable requirement for other utilities.

With respect to the analog-to-digital (A/D) and digital-to-analog (D/)A checking portion of recurrent evaluation, he was unsure. He was going to have his simulator software expert call us to confirm whether they do it and how often. He did indicate that this approach sounded more reasonable than the intermittent fidelity tests that ANS 3.5 sugg[illegible]s.

With respect to using operators to critique the simulator after they have experienced a transient, he indicated that they already do this.

The approach to "decertifying the simulator" after a plant modification until it can be properly updated seemed reasonable to him. He indicated that they have full-time simulator hardware and software experts and they are judicious about keeping the simulator up-to-date. Minor updates are made as they occur in the plant and major updates are made somewhat less frequently. However, he indicated that smaller utilities may have some difficulties in satisfying this requirement if they do not have full-time staff.

In summary, he indicated that the approach sounded reasonable and he could not think of any other tests that should be added. We do not anticipate any changes in the methodology or content of the handbook based on this contact.

CONTACT REPORT

Nuclear Power Plant Utility
8/14/84

Main Topic of Conversation: The reasonableness of our approach to simulator testing.

The conversation began with a description of the general approach we were proposing for simulator testing. Then, we went through the procedure step-by-step soliciting his comments along the way.

With respect to the fidelity evaluation portion of initial simulator testing, he felt that our "ten most critiacl displays for each task" was very reasonable and better than the approach outlined in ANS 3.5. He indicated that there was a significant amount of data which was collected from the actual plant during startup that was potentially useful. On the negative side, he was concerned that the requirement for plant-specific best-estimate models during initial simulator testing would be an undue burden. If we were willing to accept generic best-estimate models (which seems like a contradiction of terms), then this would be better.

With respect to direct operator performance measurement, he felt that conceptually this was a very good idea, particularly if we collected data on operators' opinions of the simulator during this phase. However, he was concerned that operator performance measurement would be difficult and, if not carefully monitored, could result in this portion becoming a "rubber stamp." Additionally, he felt that experienced operators with no experience in the simulator would be difficult to locate.

The concept of accepting the simulator on a task-by-task basis appealed to him.

The A/D and D/A conversion checking portion of recurrent evaluation is something that they perform every day. Apparently, all Singer simulators have this capability built-in. He perceived that the only additional effort would be the paperwork associated with documentation.

With respect to tracking plant modifications, he suggested that our procedure was perhaps too rigid. He indicated that they are receiving approximately 200 plant modifications per month and that the simulator would be constantly decertified if they did not get a time allowance for an update. Furthermore, he distinguished between modifications which affect plant dynamics, which should be addressed immediately, and modifications which affect control-room "cosmetics," which are less critical and more difficult to implement. He suggested that the modifications which fit into the latter category should only be required annually, unless the control-room changes also affect plant dynamics.

With respect to the use of operator data after plant transients, he felt that this was also a good idea. He also suggested that the use of actual plant parameter data (e.g., collected via an SPDS) would be beneficial.

He also added that we may want to compare photographs of the simulator and plant control rooms every three years as a check to ensure that all plant modifications have been included.

With respect to changing our methodology, we think that we want to include the concepts he defined with respect to plant modifications. We should differentiate between control room and plant dynamic changes and reflect this in our methodology. We did not agree with his concern about the use of BEAs, simply because some of the previous information we have obtained has been much more positive in this regard.

CONTACT REPORT

Nuclear Power Plant Contractor
8/13/84

Main Topic of Converstaion: The reasonableness of the approach proposed in the handbook.

The conversation began with a description of the general approach we were proposing for simulator testing. Then, we went through each step of the approach in greater detail and solicited his comments on the feasibility of the techniques.

With respect to fidelity evaluation, he indicated that in principle the proposed approach sounds good. He had apparently been involved with a research project which compared simulator models to BEAs. We discussed some of the statistical analyses that he performed including average difference scores, trend differences, and measures of the plant's periodic behavior. In this study, he was finding a difference of approximately 3% which he felt was very good (with which we would agree), but was not up to ANS 3.5 requirements. He agreed that BEAs were probably the best basis for comparison for many tasks, although he was skeptical about the ease of collecting these data.

With respect to the direct measurement of operator performance, he felt that it was "unreasonable and unrealistic." He based this primarily on the difficulty of measuring operator performance to determine whether they were performing inadequately, both in the simulator and plant. He had been involved in some automated operator performance measurement work previously and he was very skeptical about that technology. He also indicated that licensing exams do not always test the operator's ability to perform the task. Therefore, an operator's passing a licensing exam on the simulator should not be expected to indicate that he can perform the task in the plant.

With respect to A/D and D/A checking, he also indicated that they do this daily.

With respect to the updating process he felt that the "time-frame is critical." He also indicated that hardware changes are very difficult and should not be required more than once a year. However, plant dynamics (as represented in simulator software) can be updated more frequently. This whole discussion was very consistent with the one we had with the nuclear power plant utility on August 14.

With respect to using operators after plant transients, he felt that this could be "logistically difficult." Rescheduling of operators so that those involved in the transient can go to the simulator will be difficult. Additionally, simulators are generally booked so this may require some significant rescheduling of training. He clearly indicated that this sort of approach was desirable, but we should be somewhat cautious on placing too much value on these operator's opinions. He also indicated that the simulator could be properly configured within two weeks to represent a transient, although some simulators could never be reconfigured to simulate any transient, and this potential limitation should be reflected in our handbook.

CONTACT REPORT

Best-Estimate Engineering Modeler
8/15/84

Main Topic of Conversation: The availability of best-estimate engineering models for the testing of nuclear power plant simulators.

The conversation began with a description of the general approach we were proposing for simulator testing. Then, we focused the conversation on our reliance on best-estimate engineering models for fidelity testing.

He indicated that they had developed some quite detailed and large best-estimate engineering models. Specifically, he indicated that these models included 200-300 cells per plant, whereas some BEA models had as little as 20-30 cells. His area of modeling expertise was PWRs but he indicated that there were other portions of their laboratory that do BWR modeling.

To set up a model for a specific plant takes about one man-year. Once this initial model is developed, they can run scenarios of the plant under a variety of conditions. The run-times of the worst scenarios are about 3-4 CPU seconds per second of transient simulated. The cost of CPU time on their computer is about $800/hour.

His experience indicated that the one man-year includes all plant model data collection and, if the plant designer and utility cooperates, these data can be obtained without any great difficulty.

Furthermore, the output of their models can be put in any format desired including different sampling rates and different variables. Through the "Nuclear Plant Analyzer," they can present this information graphically.

With respect to our methodology, this conversation seemed to confirm the appropriateness of our relying on best-estimate engineering models for simulator validation. He provided further evidence that this approach is feasible, albeit somewhat expensive. We guess that we could obtain data on 30-40 scenarios for about $200-300k. Given the cost of simulators, this would only represent a small percentage (about 5%) of overall simulator cost.

CONTACT REPORT

Best-Estimate Engineering Modeler
8/16/84

Main Topic of Conver ation: The availability of best-estimate engineering models fc · the testing of nuclear power plant simulators.

The conversation began with a description of the general approach we were proposing fc simulator testing. Then, we focused the conversation on our reliance on best-estimate engineering models for fidelity testing.

His first response was "You're really on the hot seat." Apparently, he had spent last week at the Summer Computer Simulation Conference in Boston and had been speaking to several utility people aboult the use of BEAs for simulator evaluation. He said that these utility individuals indicated that they were testing the simulators largely against plant data. When he argued that they should be using BEAs for "beyond the boundary" conditions, he claimed that the utilities agreed, albeit grudgingly.

His feeling was that BEAs could and should be used for simulator comparison. He understood that many utilities had best-estimate models and most of those who did not could get them from the Electric Power Research Institute.

Within his group at the laboratory, they have been working on building best-estimate simulations that would allow for dynamic operator input. This means that, while a model of a plant event is running, they could change the model to include new crises and operator input. This would make, in essence, a best-estimate simulator, except that they do not currently run in real-time. It sounds very similar to what we saw at one of the vendors. They can develop these models for any plant, given cooperation with the utility and thirteen man-months. Once these models are developed, they can run virtually any kind of event on that plant. The cost in computer time is about $10,000-$20,000 for a large event and $2,000-$5,000 for a small event.

With respect to our methodology, this conversation again confirmed the availability of best-estimate models for simulator testing. Every contact with best-estimate modelers has confirmed this. The only skeptics seem to be the individuals at the utilities who are either unaware of the technology or are hesitant to make the needed investment.

CONTACT REPORT

Nuclear Power Plant Utility
8/21/84

Main Topic of Conversation: Reasonableness of our proposed approach to simulator testing.

The conversation began with a description of the general approach we were proposing to simulator testing. Then, we went through each step of the methodology in more detail soliciting opinions on the way.

With respect to fidelity measurement, they expressed some hesitations. They agreed that we needed to compare against one of the three reference data sources, however, they were of the opinion that training simulator models were more detailed than many best-estimate codes.

With respect to direct operator performance assessment, they said that it sounds interesting. However, there was a general lack of transients which would mitigate against this approach. Also, they thought that it would be difficulat or impossible to replicate a transient on the simulator within a few weeks.

With respect to A/D conversion checking, this is done automatically on their simulator.

With respect to simulator modification to reflect plant updates, they agreed with this "100%." They thought that 30 days was reasonable, although 90 days may be more reasonable in some cases.

With respect to simulator operators returning to the simulator after a transient, they expressed concern that they could configure the simulator.

In summary, they had many hesitations about the methodology. Some of their concerns about BEAs do not appear to be well founded. However, their concerns about replicating transients on the simulator may be more serious. Other contact have indicated that this was feasible.

NUREG/CR-4188, Vol. 2
ORNL/TM-9570/V2
Distribution Category RX

INTERNAL DISTRIBUTION

1. L. S. Abbott
2-10. R. J. Carter
11. G. F. Flanagan
12-16. P. M. Haas
17. C. C. Jorgensen
18. H. E. Knee
19. P. A. Krois
20. F. C. Maienschein
21. A. P. Malinauskas
22. J. J. Manning
23. F. R. Mynatt
24. A. Zucker
25. Central Research Library
26. Y-12 Document Ref. Section
27. Laboratory Records ORNL, RC
28. ORNL Patent Office

EXTERNAL DISTRIBUTION

29. Director, Division of Reactor Research and Technology, DOE, Washington, D. C. 20545.
30. Office of the Assistant Manager for Energy Research and Development, DOE-ORO, Oak Ridge, TN 37830.
31-33. K. R. Laughery, Micro Analysis and Design, 9132 Thunderhead Drive, Boulder, CO 80302.
34. C. R. Bovell, Physics Applications, Inc., 2023 Bice Road, Soddy-Daisy, TN 37379.

Electric Power Research Institute, 3412 Hillview Avenue, P. O. Box 10412, Palo Alto, CA 94303.

35. J. F. O'Brien
36. H. L. Parris
37. J. P. Sursock

Institute of Nuclear Power Operations, 1100 Circle 75 Parkway, Suite 1500, Atlanta, GA 30339.

38. Z. T. Page
39. K. Strahm
40. J. C. Lowry, Allen Corp. of America, 401 Wythe Street, Alexandria, VA 22314.
41. L. H. Gray, 866 Greenwood Avenue, Clarksville, TN 37040.

U.S. Nuclear Regulatory Commission, Washington, DC 20555:

Division of Risk Analysis and Operations:

42-44. J. A. Norberg

Division of Human Factors Technology:

45. B. A. Boger
46. H. R. Booher
47. D. B. Jones
48. W. T. Russell
49. J. A. Wachtel
50. D. Ziemann

51- 52. Technical Information Cente.
53-452. Given distribution as shown in NRC Category RX, Human Factors Safety (NTIS - 10).

NOTICE

When you no longer need this report, please return it to P. M. Haas, Engineering Physics and Mathematics Division, Bldg. 6025, Rm. 10W, Oak Ridge National Laboratory, P. O. Box X, Oak Ridge, Tennessee 37831. Thank you.

NRC FORM 335 (2-84) NRCM 1102, 3201, 3202

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

SEE INSTRUCTIONS ON THE REVERSE

1. REPORT NUMBER (Assigned by TIDC, add Vol. No., if any)
NUREG/CR-4188, Vol. 2
(ORNL/TM-9570/V2)

2. TITLE AND SUBTITLE
Nuclear Power Plant Simulation Facility Evaluation Methodology: Technical Bases

3. LEAVE BLANK

4. DATE REPORT COMPLETED
MONTH: November YEAR: 1985

5. AUTHOR(S)
K. R. Laughery, Jr. and R. J. Carter

6. DATE REPORT ISSUED
MONTH: January YEAR: 1986

7. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Engineering Physics & Mathematics Division
Oak Ridge National Laboratory
P. O. Box X
Oak Ridge, TN 37831

8. PROJECT/TASK/WORK UNIT NUMBER

9. FIN OR GRANT NUMBER
B0466

10. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Division of Risk Analysis & Operations
Office of Nuclear Regulatory Research
U. S. Nuclear Regulatory Commission
Washington, DC 20555

11a. TYPE OF REPORT
Final

b. PERIOD COVERED (Inclusive dates)

12. SUPPLEMENTARY NOTES

13. ABSTRACT (200 words or less)

This report is Volume 2 of a two-part document which describes a project conducted to develop a methdology to evaluate the acceptability of nuclear power plant (NPP) simulation facilities for use in the simulator-based protion of NRC's operator licensing examination.

The proposed methodology is to be utilized during two phases of the simulation facility life-cycle, initial simulator acceptance and recurrent analysis. The first phase is aimed at ensuring that the simulator provides an accurate representation of the reference NPP. There are two components of initial simulator evaluation: fidelity assessment and a direct determination of the simulation facility's adequacy for operator testing. The second phase is aimed at ensuring that the simulation facility continues to accurately represent the reference plant throughout the life of the simulator. Recurrent evaluation is comprised of three components: monitoring reference plant changes, monitoring the simulator's hardware, and examining the data from actual plant transients as they occur.

Volume 2 describes the development of and technical bases for the evaluation methodology, including a discussion of the major issues, research base, and judgements/decisions made by the research team.

14. DOCUMENT ANALYSIS – a. KEYWORDS/DESCRIPTORS

simulator
simulation facility
simulation facility evaluation
simulator examination
simulator fidelity assessment

b. IDENTIFIERS/OPEN-ENDED TERMS

15. AVAILABILITY STATEMENT
Unlimited

16. SECURITY CLASSIFICATION
(This page) Unclassified
(This report) Unclassified

17. NUMBER OF PAGES

18. PRICE